

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Jan</u>	NA Sequence (#) _____	STN _____
Searcher Phone #: <u>448</u>	AA Sequence (#) <input checked="" type="checkbox"/> _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>4/23/03</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>4/23/03</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <input checked="" type="checkbox"/> _____
Clerical Prep Time: <u>10</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>10</u>	Other _____	Other (specify) _____

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:28:09 ; Search time 50 Seconds
(without alignments)
1543.919 Million cell updates/sec

Title: US-10-034-849-2

Perfect score: 4188
Sequence: 1 MTEVFTVLDSSYEYVQKEP.....LKASATGQKTLFDFLAKSK 803

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3033.5	72.4	803	2	B56277
2	2434	58.1	784	2	E72515
3	2246	53.6	781	2	UC7382
4	1581	37.8	781	2	A69312
5	1287	30.7	764	2	S75407
6	1236.5	29.5	775	2	S35543
7	1207.5	28.8	771	2	C75023
8	1197.5	28.6	775	2	S67920
9	965.5	23.1	1312	2	S68593
10	955.5	22.8	1235	2	C71210
11	915	21.8	586	2	C69028
12	771.5	18.4	1086	2	T40242
13	768.5	18.4	1086	2	S71551
14	763.5	18.2	1086	2	T43266
15	760.5	18.2	1107	2	T43266
16	757.5	18.1	1106	1	A39299
17	751	17.9	1038	1	UC5757
18	749	17.9	1038	1	T18232
19	743.5	17.8	1105	1	S40243
20	741.5	17.7	1088	2	T05731
21	731.5	17.5	1084	1	S19661
22	718.5	17.2	901	2	E84210
23	705	16.8	1094	2	S22573
24	691.5	16.5	1097	1	RNBV3
25	671	16.0	1634	2	E64410
26	649.5	15.5	879	2	A56277
27	647.5	15.5	872	2	UC7380
28	628.5	15.0	875	2	UC5186

30	625.5	14.9	959	2	F72763	probable DNA-direc
31	598.5	14.3	882	2	S23019	DNA-directed DNA p
32	598	14.3	1081	2	T20698	hypothetical prote
33	594.5	14.2	882	2	F90201	DNA polymerase I (
34	579.5	13.8	1462	1	DJHUC	DNA-directed DNA p
35	578	13.8	1465	2	S45628	DNA-directed DNA p
36	569	13.6	787	2	E82327	DNA polymerase II
37	564.5	13.5	844	2	T31321	DNA-directed DNA p
38	527.5	12.6	787	2	G83410	DNA-directed DNA p
39	526	12.6	3122	2	T17202	DNA-directed DNA p
40	524.5	12.5	1015	1	DJBR2L	DNA-directed DNA p
41	522.5	12.5	1339	1	S20052	DNA-directed DNA p
42	517	12.3	1505	2	S28079	DNA-directed DNA p
43	515	12.3	1009	1	DJBEM2	DNA-directed DNA p
44	511	12.2	1026	2	T03108	DNA-directed DNA p
45	508.5	12.1	1513	2	T28158	probable DNA-direc

ALIGNMENTS:

RESULT 1	
B56277	
DNA-directed DNA polymerase (EC 2.7.7.7) II - Pyrodicticum occultum	
C:Species: Pyrodicticum occultum	
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #ext_change 20-Jun-2000	
C:Accession: B56277	
R:Uemori, T.; Ishino, Y.; Doi, H.; Kato, I.	
J. Bacteriol. 177, 2164-2177, 1995	
A:Title: The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-like DNA poly	
A:Reference number: A56277, PMID:95238250; PMID:7721707	
A:Accession: B56277	
A:Status: Preliminary	
A:Molecule type: DNA	
A:Residues: 1-803 <DEM>	
A:Cross-references: GB:D38574; NID:9807829; PIDN:BA07580.1; PID:9807830	
C:Superfamily: herpesvirus DNA-directed DNA polymerase	
C:Keywords: nucleotidyltransferase	
Query Match 72.4%; Score 3033.5; DB 2; Length 803;	
Best local Similarity 70.9%; Pred. No. 1.4e-176;	
Matches 571; Conservative 106; Mismatches 121; Indels 7; Gaps 5;	
Qy	1 MTEVFTVLDSSYEYVQKEPQVITINGIANGERSVILDSFPYFALAPGADP--KO 57
Db	1 MTEVFTVLDSSYEYVQKEPQVITINGIANGERSVILDSFPYFALAPGADP--KO 57
Qy	58 VAQIRALSPKSPITIGVEDDKRKYRGRPRVLRIRTVLPEAVREYRELYKQNDGVEDVL 117
Db	61 IASIRLSPKSPITIGVEDDKRKYRGRPRVLRIRTVLPEAVREYRELYKQNDGVEDVL 120
Qy	118 EADIRFAMRYLIDHDLFPPTVRYVEAEPLLENKMGFRVYKQVYLVKSPREPLYGALAPTKL 177
Db	121 EADIRFAMRYLIDHDLFPPTVRYVEAEPLLENKMGFRVYKQVYLVKSPREPLYGALAPTKL 177
Qy	178 PDRLAFDIEVYSKQSPPEPDPVIVIAVKTDDGDEVLFIEGKDDKRRIFEEVYEVK 237
Db	179 PDRLAFDIEVYSKQSPPEPDPVIVIAVKTDDGDEVLFIEGKDDKRRIFEEVYEVK 237
Qy	238 RYDDDIIVGNNHFPDMPVYLRRARITGILQVTRRGAAPTSSVGHVSVPRLNVDLY 297
Db	239 RYDDDIIVGNNHFPDMPVYLRRARITGILQVTRRGAAPTSSVGHVSVPRLNVDLY 297
Qy	298 DYAEEMPEIKIKLSEVAEYLVGKKSERYIIMWEIIPDYDDPKRPLLLQYARDVRA 357
Db	299 DYAEEMPEIKIKLSEVAEYLVGKKSERYIIMWEIIPDYDDPKRPLLLQYARDVRA 357
Qy	358 TYGIAELTLPFPAIQSVYTLPLDQVANSVGRLEWYILRAAFKMKELVNVREPEET 417
Db	359 TYGIAELTLPFPAIQSVYTLPLDQVANSVGRLEWYILRAAFKMKELVNVREPEET 417
Qy	418 YRGATVLEPGRGVHNAVLDPSFSSMYPNIMIKYNNVGPDTLVRRGKCGEC--GCMGAPVVK 476
Db	418 YRGATVLEPGRGVHNAVLDPSFSSMYPNIMIKYNNVGPDTLVRRGKCGEC--GCMGAPVVK 476

Db 419 YGAVVLPKLGHEVNVVLDFFSSMPSIMIKVNGPDTIVDDSECECKYGGCYVAPEVG 478
 Qy 477 HFRRCPPGPFKTVLERLELRKRVBAEMKKYPPDSEYRLLDEROKALKVLNANSYGYM 536
 Db 479 HFRRCPPGPFKTVLERLELRKRVBAEMKKYPPDSEYRLLDEROKALKVLNANSYGYM 538
 Qy 537 GMSGARVYCRECAKAVTAMGRHLIRTAIRNARKGLKVIYGDTSLSFVYDPEKVENFLK 596
 Db 539 GMSHARVYCRKCAAVTAMGRHLIRTAIRNARKGLKVIYGDTSLSFVYDPEKVENFLK 598
 Qy 597 IIKELGELTLEKRVNKLFFPTAKKRVAGLLEDRIDIVGFPAVNGDNCCLAKVQTKV 656
 Db 599 FVEKLGPEIKIDIKYKVFTEAKKRVYGLLEDRIDIVGFPAVNGDNCCLAKVQTKV 658
 Qy 657 VEIVTKSEVNAKAVEYRKIVKELEKVPLEKVIWKTLSKRLSEYTTAEAPHVAARQM 716
 Db 659 AEIVLTNGVNDKALISYIREVIAKQREGKVPITKLIWKTLSKRLSEYTTAEAPHVAARQM 718
 Qy 717 LSAGYRVSPGDKIGYVIVKGGRIISORAMPYFVNDPSQIDVTVYVDHQTIPALRLIGY 776
 Db 719 KEAGYEVSPGDKIGYVIVKGGSGSVSRAPYFVNDPSQIDVTVYVDHQTIPALRLISY 777
 Qy 777 FGITEKLLKASATGQKTLFDFPLAKK 801
 Db 778 FGITEKOLKAAATVQSLFDFPFASK 802

RESULT 2

probable DNA-directed DNA polymerase APE2098 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: E72515
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah
 ame, H.; Takamaya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Reference number: A72450; M0ID:99310339; PMID:10382966
 A:Accession: E72515
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-784 <KAM>
 A:Cross-references: DDBJ:AB000663; NID:95105654; PIDN:BAAB1109.1; PID:G5105797
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE2098
 C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 58.1%; Score 2434; DB 2; Length 784;
 Best Local Similarity 59.1%; Pred. No. 3,76-140;
 Matches 469; Conservative 113; Mismatches 189; Indels 22; Gaps 5;

Qy 17 GCEPOVITMGIAENGERVYLIDSRFPYFALAFAGADPKQVACORIRALSRRPSPIGYE 76
 Db 3 GSTVITLMGAGDSKSHVVFGEFRPIFYVLPDSYGLDQLAAMIRLSRPSPIYS 62
 Qy 77 DDKKRYGRRPRVLRIRTVLPEAVREYRELKVNVDYEDVLEADIRFARKYLIIDHDLPEF 136
 Db 63 RVRREFIGREVALKVTTLVPAVSREYRAVARLOGVADVLEADIRFARFIIDIRLYPM 122
 Qy 137 TWYVEAEPLLENKMGFRVDKVVYL-----KSREPLYGEALAPTKLPDLRLIADI 187
 Db 123 RUYVAEVAEVAVPHGYSDVRAVYLSGDIRBEDTRIOEDPFLKG-----LRVARDI 172
 Qy 188 EYVSKGSPREPRDPVIVIAVKTDGDEVLFIAEGDKDRKPIREFEYVYKRVDPDIYGY 247
 Db 173 EYVSKMRTPPDKOPVIMIGLQQAQGEIRILEAEDRSKVIAGFVERKYSIDPVIYGY 232
 Qy 248 NNNHEDMPYLLRRARILIGILDTYRVGAEPSTSHGVSPGLINVDLYDAEEMPEIK 307
 Db 233 NQNRDMPYLLVERARVIGVLANGR-SVEPQGLYGHYSVSGRLNVDLIDRAELHAYK 291
 Qy 308 IKSLEVAEYVIGVWKKSERVIINMEIPDYNDPKKRLILQIARDDVRAITYGLAEKILP 367

Db 292 VKTLEEVADYGVVIGERVYLTLEWQIGEWDDPCKREILKRYLRDDVSTWGLAKFIP 351
 Qy 368 FATQLSYTMGRLPDQVAMSVGRLEWYVILRAAFMKELVNRVERDEEYRATYLEPL 427
 Db 352 FGAEISQVSGPLPDQVMAASVGRLEKRLIRBAALGELVNRVERDEEYRATYLRK 411
 Qy 428 RGHENIATLDPSSYPMIMIKVNGPDTLVPRGEKCECCMAPEYKRRFRCPPGF 487
 Db 412 PGHEIATLDPASVPMIMIKVNGPDTLVPRGEYEEBEYTAPEYKRRFRCPPGF 471
 Qy 488 KTVLEELLRKRVBAEMKKYPPDSEYRLLDEROKALKVLNANSYGYMGSARVYCE 547
 Db 472 KTLERFLSMRQRISEKKGPPDSEYRLLDEROKALKVLNANSYGYMGPAPARVYCRE 531
 Qy 548 CAAVATAMGRHLIRTAIRNARKGLKVIYGDTSLSFVYDPEKVENFLIKIKELEGEIK 607
 Db 532 CAEAVTAMGRSIIIRTAIRKAGEIGLEVIYGDTSLSFVNDPEKVERLIRFEVELEFDIK 591
 Qy 608 LEKVYRLFFTEAKKRVAGLLEDRIDIVGFPAVNGDNCCLAKVQTKVVEIVLKTSEVN 667
 Db 592 VDKVYRVVFFTEAKKRVYGLTVQKIDIVGFPAVNGDMSLAKETQKVAEIVLKTGSD 651
 Qy 668 KAVEYRKIVKELEKVPLEKVIWKTLSKRLSEYTTAEAPHVAARQMISAGYRVSPG 727
 Db 652 EAVDYRNRIEKLARGVDNRKVIWKTILTRPSMYEARQHVYALMLMERAGIKVEPBA 711
 Qy 728 KIGYVIVKGGRIISORAMPYFVNDPSQIDVTVYVDHQTIPALRLIGYFITEKTLAS 787
 Db 712 KIGYVIVKGGSGSVSRAPYFVNDPSQIDVTVYVDHQTIPALRLIYGVTEKRLKG 770
 Qy 788 ATGQKTLFDFPLAKK 800
 Db 771 GR-QSTLIDPFMR 782

RESULT 3

JC7382
 DNA-directed DNA polymerase (EC 2.7.7.7) B3 - Sulfuriophthora ohwakuenis
 N:Alternate names: DNA polymerase alpha, beta, gamma; DNA polymerase I, II, III
 C:Species: Sulfuriophthora ohwakuenis
 C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
 C:Accession: JC7382
 R:Iwai, T.; Kurosawa, N.; Itoh, Y.H.; Kimura, N.; Horikuchi, T.
 DNA Res. 7, 243-251, 2000
 A:Title: Sequence analysis of three family B DNA polymerases from the thermophilic
 A:Reference number: JC7380
 A:Accession: JC7382
 A:Molecule type: DNA
 A:Residues: 1-781 <IMA>
 A:Cross-references: DDBJ:AB032376
 C:Comment: This enzyme has both 3'-5' exonuclease and polymerase activities, and plays
 C:Genetics:
 A:Gene: B3
 C:Keywords: DNA replication; exonuclease; metal binding; nucleotidyltransferase

Query Match 53.6%; Score 2246; DB 2; Length 781;
 Best Local Similarity 54.3%; Pred. No. 9,96-129;
 Matches 432; Conservative 141; Mismatches 203; Indels 20; Gaps 10;

Qy 6 FTVLDSSYEVNGKEPOVITMGIAENGERVYLIDSRFPYFALAFAGADPKQVACORIRAL 65
 Db 5 FTLDPSYDVVENKRVYIYIWDKGNRVLLKKFRFYFALVDSDYINIDIRKISLTL 64
 Qy 66 SRKPSIIGVEDDKRYGRRPRVLRIRTVLPEAVREYRELKVNVDYEDVLEADIRFAM 125
 Db 65 SKYSPISISIDEEKYFGSPVVLKIEIVYIRAYRVYRDVAIKIKVKSVEADIRFIM 124
 Qy 126 RYLIIDHDLFPPTWYRVBAEPLKMGFRVDKYYLYKSPPELYGEALAPTKLPDLRLIAD 185
 Db 125 RISIDINLFPFWBAEVEIKEN-NFRVKKVYELK-KINKLYED-----KIPELKVLAF 177
 Qy 186 DIEVSKGSPREPRDPVIVIAVKTDGDEVLFIAEGDKDRKPIREFEYVYKRVDPDIIV 245

Query Match 30.74; Score 1287; DB 2; Length 764;
 Best Local Similarity 34.84; Pred. No. 1.9e-70;
 Matches 280; Conservative 167; Mismatches 305; Indels 52; Gaps 16;

QY 6 FTVLDSYEVVNGKEPOVIIMGIAENGERVVLIDRSFRPYFALLAPGADKPQVAAQRIRAL 65
 DB 5 FTFLDSYEVVNGKEPOVIIMGIAENGERVVLIDRSFRPYFALLAPGADKPQVAAQRIRAL 59
 QY 66 SRKSPPIIGVEDDKRYKFGFRPRVLRITVLEAVREYRELKVNQVDEVDLEADIRFAM 125
 DB 60 -KNGEALQITKVKRKYGINIVDALLIGTSTQIKKREKISLNNIKIPADIRYTM 118
 QY 126 RYLIDHDLPEFTWYRVEAPLENMGFRVDKYLKSRPPELYGEALAPTKLPDLRIAP 185
 DB 119 RYSLDPEDLRPFMTWFAEYNEVKFD-GFRTRKAYIIDKLISHYEG-----NMDELKTIGV 171
 QY 186 DIEVYSKQSPRPERDPIYIAVKTDDGEVLFTAEKGDPRKPIREVEYVKRPDPIIV 245
 DB 172 DQGIYSKQSGLNPRKDPVIMVMSLMSKEGMPQSLDEGIDDKIRRVVDIILAYDPIIF 231
 QY 246 GYNNHFMFMPYLBARATIGIKLDVTRRGAEPTTSVGHVSFGRLANDLVYAEEME 305
 DB 232 VIDSDLPKKTITTEBASLGKIDIGKIGSEVSGTIGHYSISGRNLVDTGLVNERG 291
 QY 306 IKIKLSEVAYELGVMMKSERVIINMWEIPDYDDPKRPLLOYADVRAITYGLAEKI 365
 DB 292 LGHVLDLIDVSNLTGI--SPSRYSFKWEISRWDNEKNRRIRIERYSIENRSTIYLLGNYL 349
 QY 366 LPPALQLSYVTGLPDDOVGAMSVGFLEWYLIRAAFKMKELVPRVERPE--ETYGALIV 423
 DB 350 LSTSELVYIVGLPLDKLSVASWGNRIETSLIRITAKSGELIPTRMDNPRPKIKKNII 409
 QY 424 LEPRLGVHENIAVLDFSSMYPNIMIKYNVGPDTLVPRPKGCEGCEAEAVGHRPRCP 483
 DB 410 IOPKGIYTDVYVLDISSVYSLVRKENIAPDLVK--EQCDD--CYSSPISMYKRRER 465
 QY 484 PGFFKTVLERLLELRKVRRAEMKYPSPDSPEYRLDEROKALKYLANASGYMGWGAM 543
 DB 466 SGLYTKFLDELSNVR-----DSNKIKYVIE-----LISFNDYHVMWNAAM 506
 QY 544 YGCECAKAVTAMGRLIRTAIINIAARKLGKVIYGPDTSLFYT-YDPRKYNFKIKIEL 602
 DB 507 YSREIASAFDEFSNEIIRFIIDLKSSGDLVLANDLIFVTTGSRKKNBELTKINSLY 566
 QY 603 GPEIKLEKVKYKLPFTTAKKRYAGLEDGRIDIVGFVARD---WCEIAKVVQRYVE 658
 DB 567 NIDVAKIIFYKSLVLD--NNRIYAGLSBGKIDI---ARKGEEDMNLCEIARNIKRIIE 621
 QY 659 IYUKTSEVNAKAVEYRKIVLEBEGKVPTEKVIYMKULSKLSEYTTBADPHVAARML 718
 DB 622 ETLISKDYKAKIKLVKSVIKLRGEPDEBELITAKERLDINMYNNQLPFFVAAKALQ 681
 QY 719 AGRYVSPDGKIGYIVVKGGRISORAMPYFMVPSQIDVTYYVHQIIPALRIIYGF 778
 DB 682 SGLYSKOSKIGYIVVKGGLPNDRAEPFLVEKRRIDIEYVD-QIPRETIKLKPLG 740
 QY 779 ITEKKL-KASATGOKTLFDFLAKK 801
 DB 741 VNEESLKKTNITDIDLDFGASKKK 764

RESULT 6
 S35543
 DNA-directed DNA polymerase (BC 2.7.7.7) - Pyrococcus furiosus
 C:Species: Pyrococcus furiosus
 C:Date: 31-Dec-1993 sequence_revision 02-Aug-1994 #text_change 20-Jun-2000
 C:Accession: S35543; S44596
 R:Unprot: T; Ichino, Y.; Toh, H.; Aada, K.; Kato, I.
 Nucleic Acids Res 21, 259-265, 1993
 A:Title: Organization and nucleotide sequence of the DNA polymerase gene from the archae
 A:Reference number: S35543; M01D:93181200; PMID:8441634
 A:Accession: S35543

A:Molecule type: DNA
 A:Residues: 1-775 <UEM1>
 A:Cross-references: EMBL:D12983; NID:g216917; PIDN:BAA02362.1; PID:g216918
 A:Accession: S44596
 A:Molecule type: protein
 A:Residues: 1-12 <UEM2>
 A:Gene: pol
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; nucleotidyltransferase
 F.1.775/Product: DNA-directed DNA polymerase #status experimental <MAT>

Query Match 29.54; Score 1236.5; DB 2; Length 775;
 Best Local Similarity 36.24; Pred. No. 2.3e-67;
 Matches 301; Conservative 153; Mismatches 283; Indels 95; Gaps 23;

QY 8 VLDSSYEVVNGKEPOVIIMGIAENGERVVLIDRSFRPYFALLAPGADKPQV---AQRIR 63
 DB 2 ILVDYITEGKFPVIRLK-KENKRIKHDTIRPIYLLAPDSDKIEVKITGERKG 60
 QY 64 ALSRPKSPITIGVEDDKRYKFGFRPRVLRITVLEAVREYRELKVNQVDEVDLEADIRF 123
 DB 61 KIVR---IVDEKVEKKEFKGKPTWVKULSHPDQVPTIREKVRPAVVDIFEYDIF 116
 QY 124 AMRYLIDHDLPEFTWYRVEAPLENMGFRVDKYLKSRPPELYGEALAPTKLPDLRI 183
 DB 117 AKRYLIDKGLI-----PMEGE-----BELKITL 138
 QY 184 APDIEVYSKQSPRPERDPIYIAVKTDDGEVL-----FTAEKGDPRKPIREVEY 235
 DB 139 APDIEVYSKQSPRPERDPIYIAVKTDDGEVL-----FTAEKGDPRKPIREVEY 235
 QY 236 VKRYDPIIVGNNHFMFMPYLBARATIGIKLDVTRRGAEPTTSVGHG--VSYPGL 292
 DB 198 IREKDDPIIVTNGSPDFPLAKRAKGIKLTIGRD-GSEPMQRIQDMTAVYKRI 256
 QY 293 NVLDVYAEEMPEIKISLEEVAEYGVMMKSERVIINMWEIPDYDDPKRPLLOYAR 352
 DB 257 HPDLVHIVIRTNIPFTYVLEAVEAL-FGKRPKRYAD--EIAAMSGENLVRVAKSM 313
 QY 353 DVPATYGLAEKILPRLQLSYVTGLPDDOVGAMSVGFLEWYLIRAAFKMKELVPRVER 412
 DB 314 EDKATYELGKFLPMEIQSLVGPRLMDRSSTGVLWFLKALYRNENAVNPKS 373
 QY 413 RPE-----ETYGALIVLEPRGVHENIAVLDFSSMYPNIMIKYNVGPDTLVPRPKGCE 466
 DB 374 EBYQRLREESTGTFVKEPEKGLMENIYVLDPRALYPSIIITHNVPDTLNEG----- 428
 QY 467 CGCWE-APENVKRRFRRCPPGFKTVLERLLELRKVRRAEMKYPSPDSPEYRLDEROKAL 525
 DB 429 CKNYDIAIPQVGHKFCQDIPGFIPLGHLLEEROKIKTKMKE-TODIEKILLDYRQKI 487
 QY 526 KYLANASGYMGSGARWYCECAKAVTAMGRLIRTA-INIAKGLDKVIYGDITSLFV 584
 DB 488 KLANSPFYGYVYAKARWYCECAESVTAMGKRYIELWKELEKFEFKALYIDTGLA 547
 QY 585 TYD-----PEKVENFKIKIEELG--FEIKLEKVKYKLPFTTAKKRYAGLEDGRIDI 635
 DB 548 TIPGSESEIKKKALFVKYINSLPGLLELEVEGFRGFFV-TKRRAVIDEESKVT 606
 QY 636 VGFEAVRGMCELAKEVQTKVVEIVLKTSEVNAKAVEYRKIVKLEBEGKVPTEKVIYMT 695
 DB 607 RGLIYIRRMSSIAKETQARVLTILKHGDVEAVRIYKVIQCLAYEIPPEKLIYQ 666
 QY 696 LSKRLSEYTTTAPHVAARMLNSGRVSPDGKIGYIVVKGGRISORAMPYFMV--DP 753
 DB 667 ITPPLLEIRAIQPHVAARMLNSGRVSPDGKIGYIVVKGGRISORAMPYFMV--DP 753
 QY 754 S--QIDVTYYVHQIIPALRIIYGFRTGITEKKULKASATGOKTLFDFL-AKKS 802
 DB 724 KKHAYDAEYIENQVLPALVITILEGFRKEDLRYOKTRQVGLTSMINIKKS 775

RESULT 7

DNA polymerase I PAB1128 - *Pyrococcus abyssi* (strain Orsay)

C:\species: Pyrococcus abyssi

C:\Date: 20-Aug-1999 #sequence 20-Aug-1999 #text_change 20-Jun-2000

C;Accession: C75023

R; anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stru

A;Reference number: A75001

A;Accession: C75023

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-771 <KAW>

A;Cross-references: GB:AJ248288; GB:AL096836; NID:q5458960; PIDN:CAB50625.1; PID:q545913

A; Experimental source: strain Orsay

C;Genetics:

A;Gene: polI; PAB1128

Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match	Score	DB 2;	Length
28.8%;	1207.5;		771;
Best Local Similarity	35.7%;		
Best Local Similarity	Prod No. 130.65.		

```

QY 8 VLDSSYEVAWGKRPQVIMGIAENGERVYLIDRSFPRPYFALLAPGADPRQV----AQRIR 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 IIDADYITEDGKPIIRIFK-KEGGEFKVEYDRTPRPYITALLKODSDAIDEVKKIAENHG 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 ALSRPKSPICGEDDKKTKFGPRPRVLRIRITVLPEANVEYRELVKNOGDVEVDLEADIRF 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 KIVR----ITTEKQKCKFLGRPIEWKLYLEHPDVPARIREKHENPAVVDIPEYDIPF 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 AMRYLIDHDLFPFTVRYVAEAPLENKMGFRVDKVLYKSRPRPLVGEALAPFKLDLRL 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 AKRYLIDKGL-----TPMEGN-----BELFTL 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 APDIEVSKQSGPRPERDPVIVIAVKTDDGDLV-----FIAEGDKDRKPIREFVEY 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 139 AVDIETLTHEGE-EEGKGPIIMISYADEGAKVIMTKSIDLPEVEVSSEREMIRLVKY 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 VKRVDPDIIIVGNNNHFPMPYLLIRBARILIGTLDVTRVGAEPPTSVMH--VSPGRL 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 198 IREKOPDVIITYNGNPFPPYLLIKRAEKIGTLPGRD-NSEPKQWOGDSLAVEIKSRI 256
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 NVDLDAVABEMPEIKISLAEVAEYLGVWKKSERVIIMMWEIJDVWDPDKRPILLQYAR 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 HFDPFVPIRRTINLPTTYLAEVYEA-PEKSKKKYAH--ETAEAMEGKGIERAKTSM 313
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 DDVRATYGLAEKILDPAIQLSYVTGLPLDOVGAMSGVFLYEWYLIRAAFKMEKELPNVE 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 314 EDAKVTFELGKEFPFMEQOLARLVGOFPWDVSRSTGNLVENFLRKAYERNELAPNPD 373
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 413 RPE-----ETRGALIVLEPLRGVHENIAVLDPSSMYPRIMIKIVNGPDTLVRPEKGE 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 374 EREYERRLRESYEGYVKEPEKGLMEGIVSLDPRSLYPSIIITHNVSPDTLNR--ENCKE 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 CGCWEAPEVKHFRCRCPGCFKTVLERLLELTKRAVEAKKXPRPDSPEXRLIDEROKLX 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 432 YDV-APQVGRHFCDFPGFISLGLNGLLEEKOKIKKKRKE-SKDPVEKGLADYRQRIK 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 527 VLAANSYGYMGSGARWYCRECAKAVTAWGRHLIRTAINTARKGLKVIYGDPTSLFVY 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 ILANSYGYGYAKARWYCKEACASVTAWGRQYIDLVRRELSRPFKVLIDYTDGLVANTI 548
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 587 D-----PEKVENIKIKIIEKELG--FEIKLEKYKRLPFTAKKRYAGLLEDGIDIVG 637
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 549 PGAGHEEIKERKLKVEVEYINSLPGULBLEYEGFYARGFV-TKKRYALIDEGKIVIRG 607
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 638 FEAVRGDMCELAEVOTKVEIVLTKTSEVNKAVEYVRKIVKELESGKVPTEGLVIMTKTS 697
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 608 LEIVARMSSEIKENQAKYLEAILKHGNDVAVKLVKETEGLSKYEIRPREGLVITYEQT 667
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 638 KRLBEYTTBAPVVAAKRMLSAGIRVSPEDKIGYVIVKGCGRIISQRAMPDYFWK--DPS- 754

```

1. The first group of respondents (n = 10) was composed of students who had completed the course and were currently employed in a related field. 2. The second group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 3. The third group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 4. The fourth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 5. The fifth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 6. The sixth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 7. The seventh group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 8. The eighth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 9. The ninth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 10. The tenth group (n = 10) was composed of students who had completed the course and were currently employed in a related field.

Qy 755 - QIDVTYVDHQIIPAL

Db 725 HKYDA

RESULT 8

S67920

DNA-directed DNA po

C:Species: Thermococcus

C:\Date: 17-Feb-1998 #88

```
C:\Access: 1, Nov 1990 frequency_revision 15-Mar-1990 freql_change 18-Jun-1990
```

C:\ACCESS101: 30/1/20
B:Southworth. M.S.: Kong H.: Kucer

K, BOCHMORCH, M.S.; KOING, H.; KUCERA, R.B.; WALE, D.;
submitted to the EMBL Data Library, January 1996

A: Description: Cloning expression and modulation of the EMBL data library, January 1995

Query Match	28.6%	Score	1197.5	DB	2	Length	775
Best Local Similarity	36.1%	Pred. No.	5.3e-65				
Matches	301	Conservative	139	Mismatches	296	Indels	97
						Gaps	24

```

0Y 8 VLDSYEVVNGEPOVIIMGLAENGERVVLDRSPFFYALLAPGA---DPKOV-AORIR 63
Db 2 ILDDDYITTENGKPIVIRVK-KENGEFKIEYDRTEEPFYALLKSDSAIEDVKYKATARHG 60
0Y 64 ALSRPKSPIIIVEDDKRRYKFRPRVLRIRTVLFEAREYVELKANDGVEDVLEADIRF 123
Db 61 TVVKYKR---AEKVQKKFPLGRPIEWKLYPNHQQDPAIRDRIRAPAVVIDIYEIPIF 116
0Y 124 AMRYLIDHDLFPFTWYRVEAEPLNKKMGFRVDKYLVKSRBPLYGALAPTKLPDRIL 183
Db 117 AKRYLIDKGLI-----PWEGD-----EELTML 138
0Y 184 AFDEIVYSKQSPRPREDPVIIVIAKTDGDEVLFIABGD-----RKPTEBPV 233
Db 139 AFIDETLYHEGE-BFGGPIIMISYA---DGEARVITWKIDLEPYDVVSTEEKMIRFL 195
0Y 234 EYVRKYRDPDIIVGNNNNHFDMPVLYLRARILGIDLADTVRRGAEPTVSQH---VSVPG 290
Db 196 RYVKEKDPVLITNGNFDPAVLKKRCCEBLGIFTLGRD-GSPFKIQRMGDRRAVEYKG 254
0Y 291 RLAVDLDYVAEEMPEIKIKSLIEVAEYLVGWKKSERYIIMWELPDYWDPKKRPILLQY 350
Db 255 RIHFDLPVIRIRINTLPYTLLEAVEAV-FGKPRKEKYAE--ETIAQWESQEGLEVAR 311
0Y 351 ARDVRATYGLAEKILPAPAIOLSVYTGPLDOVGAMVGRPLEYVYLRAAFKMEIYPNR 410
Db 312 SMEAKTYTELGRFPFPMEOALSRLIGOSLMDVRSSTGLVEMFLRKAYKRNELAPNK 371
0Y 411 VERPEET-----YGAIVLEBPLRGVHENIAVLDPSSMYPNIMIKYNGPDLVAPGKCG 465
Db 372 PDERELARRRGYAGVYKPEBERGLMDNIYVLDPRSLIYPSIIITHVNSPTLNE----- 426
0Y 466 ECGGCE--ABEVKRRRCRCPGPFKTVLERLLELRKRVRAEMKKYPPDSEBRLDREQ 522
Db 427 --GKEVDVAEVEGHKCKCPGFIPLIDGLLEEROKIRKMKMA-TVDPLEKKLIDYRQ 483
0Y 523 KALVLANASGYVMGSGARVYCRECAKAVTAMGRHLIRAI-NIARKGLIKVLYGDTDS 581
Db 484 RAIIILANSFYGYGYAKAKARYCKECSASTAMGREYIEMVIRLEKRFSGFYIADTDG 543
0Y 582 LFTVY---DPE--KVENPIKIIKEELG--FEIKLEKYVKRLFTEAKRYAGLIDGR 632
Db 544 LHAITPGADLWETVKKKKKEFLKYNIPKLPGLLELEYGTFYVRGPFV-TKKYVAVIDEGK 602

```

QY 633 IDIVGEAVRGDMCELAKEVQTKVVEIVLKTSEVNAKAVEYRKIVKELEBQKVPDEKVI 692
 Db 603 ITRRGLEIVRDMSEIAKEVTOARLEILKHGDVEBAVRIYKEVTEKLSKYEPPEKVI 662
 QY 693 WKTLSRLEEYTTAPPHVAAKRMLSGAVRSPGDKIGYIVVKGSGRISORAMPYEMVD 752
 Db 663 HEQITRDLRYKATGPVNAVAKRLAARGVIRGTIVISYIVKSGRIGRAIPAEF-D 721
 QY 753 PS--QIDVTVYVDHQIIPALRLIGVFGITEKKLKAASATGQKTLFDFLAKSK 803
 Db 722 PTKHRYDAEYIENQVLPAYERILKAGYRKEDRYOKTKQVGLGAWLKVKKG 774

RESULT 9

566593

DNA-directed DNA polymerase (EC 2.7.7.7) Vent, intein containing precursor - Pyrococcus
 N:Contains: DNA endonuclease (EC 3.1.-.-) PI-I; DNA-directed DNA polymerase (EC 2.7.7.7)
 C:Species: Pyrococcus sp.
 Date: 24-Aug-1996 #sequence: revision 01-Nov-1996 #text: change 24-Sep-1999

Accession: S68593
 A:Name: Southworth, M.W.; Merzha, F.B.; Hornstra, L.J.; Perler, F.B.
 Submitted to the EMBL Data Library, August 1993
 A:Description: In vitro protein splicing of purified precursor and the identification of
 A:Reference number: S68593

A:Accession: S68593
 A:Molecule type: DNA
 A:Residues: 1-1312 <XUA>
 A:Cross-references: EMBL:U00707; NID:9436492; PID:AAA67130.1; PID:9825735
 R:Yu, M.Q.; Southworth, M.W.; Merzha, F.B.; Hornstra, L.J.; Perler, F.B.
 Cell 75, 1371-1377, 1993

A:Title: In vitro protein splicing of purified precursor and the identification of a bra
 A:Reference number: S68581; MUID:94094330; -PMID:8269515
 A:Contents: annotation

R:Yu, M.Q.; Comb, D.G.; Paulus, H.; Noren, C.J.; Shao, Y.; Perler, F.B.
 EMBO J. 13, 5517-5522, 1994
 A:Title: Protein splicing: an analysis of the branched intermediate and its resolution b
 A:Reference number: S52065; MUID:95080235; PMID:7988548

A:Contents: annotation; self-splicing mechanism
 C:Function: <NUCL>
 A:Description: nucleotidyltransferase
 A:Note: DNA-directed DNA polymerase Vent
 C:Function: <ENDO>
 A:Description: endonuclease; hydrolase

A:Note: DNA endonuclease PI-PePI
 C:Superfamily: hypothetical protein PH0202
 C:Keywords: DNA replication; endonuclease; hydrolase; nucleotidyltransferase; protein sp
 11-491,1030-1312/Product: DNA-directed DNA polymerase Vent #status predicted <MT1>
 492-1029/Product: DNA endonuclease PI-I (pol Vent intein 1) #status predicted <MT2>
 F:1030-1312/Domain: DNA-directed DNA polymerase Vent extein 2 #status predicted <XT2>
 F:491-1030/Cross-link: peptide (Ala-Ser) #status experimental

Query Match 23.1%; Score 965.5; DB 2; Length 1312;
 Best Local Similarity 21.9%; Pred. No. 1.4e-50;
 Matches 300; Conservative 154; Mismatches 282; Indels 635; Gaps 23;

QY 8 VLDSSYEVVKEPVIYIWIENGERVYVLDSPFRYPYFALLAGADPKOV---AQRIR 63
 Db 2 IIDADYITDEGKPIRIK-KENGEFVEYEDNFRPYIYALLKXDSQIDVRKTIKERRG 60
 QY 64 ALSRPSPITIGVEDDKRYKGRPRVIRITVLPKAVREYELVKNVNDGVEDLEADIF 123
 Db 61 KIVR-----IIDAEKRRKFFLDGRPIEVKRLTFEPDVPARKIKRHSAYVIDIFEDIF 116
 QY 124 ANRYLIDHDLFFTWYRVEAEPLNKGFRVVDKVIYVSRPELYGSLAPTKLPDLRL 183
 Db 117 AKRYLIDKGLI-----PHEGD-----EELKL 138
 QY 184 APDIEVYSKQSPREPRDPVIVAKTDDGVL-----FLAEGDKRDKIRFVEY 235
 Db 139 APDIEITLVHEGE-BPAKGPITIMISYADEBAKVIWTKIDLPYAEVSSREMIKRFIV 197

QY 236 VKRYPDIIIVGNNHNPMPYLLRRARILGILKIDVTRVGAEPPTSVHGH---VSPGR 292
 Db 198 IREKDDPVIITNGSDPDLPTVYRAEKIGIKLPIGRD-GSEPKORLGMTVAEIKGR 256
 QY 293 NVLDYVAEMPEIKIKSLSEVAEYLVGMKSERVIMMWEIPYMPDKPKRPLLIQV 352
 Db 257 HFDLYHVIRRTINLPYTLAEVYEA-FCGKPKYVAH-EIAEMETGGLERVAKYM 313
 QY 353 DVPRATYGLAEKILFPALYLSYVGLPDDVGAMSGVGFLEMYLIRAPFKELVPRVE 412
 Db 314 EDAKVYTELGRFFPEMAQSLVGLQPLMDVSSSTGVLVEMYLLRKAERNELANKPD 373
 QY 413 RPE-----ETRYGALVLEPLRGVHEIVAVDPSSMYPNIMIKVNGPDTLVPRGSGE 466
 Db 374 EREYERLSEYAGGVYEPKGLMEGLVSLDPRSLYSIIITINVSPTLNE----- 427
 QY 467 CGGME---APYKFRFRCPGPFKYVLERLLELKKRVAEMKYPDPDSPEYRLDROK 523
 Db 428 -GREDYVAPBVGHRFCDFPGFLPSLKLDEROETIKRQKA-SKDPLEKMDLYROR 485
 QY 524 ALKYLVA----- 529
 Db 486 AIKILANSILPEBWPVLKNGKVIPIIGDPVDGLMKANQKVKXGTDTEVLEVAIGHAF 545
 QY 530 ----- 529
 Db 546 SPDRSKKARVMAVAVIRHRSGNVYRIVNSGRKITTEGHSLPYRNGDLVEATGED 605
 QY 530 ----- 529
 Db 606 VKIGDILAVPRSVNLPEKREINIVELLNLSPEETEDILITPYKGRKNFPGMLRTL 665
 QY 530 ----- 529
 Db 666 WIRGEKVRVTSRYRLRLNBLGIRLKIYDIDKSGLEKRYTLEKLVDDVVRVNGK 725
 QY 530 ----- 529
 Db 726 RELVFNPAVRVIVLMEELKEMRLGTRNGFRMGTFVIDEDPAKLGIYVSEGSARK 785
 QY 530 ----- 529
 Db 786 WKQGTGMSYVRLYNENDEVLDMELAKKPFQKVGKNYVEIPKKMAYIIFESLGGT 845
 QY 530 ----- 529
 Db 846 LAENKEVPVIFTSSKGVNAFLGYFTGGDVHPSKVRSLSTKSELLVGLVLNLSIG 905
 QY 530 ----- 529
 Db 906 VSAIKLGDGSGVRYVNEBELKFTERYKKKNVYSHIVKDLKETFGKVFQKNISYKFF 965
 QY 530 ----- 529
 Db 966 RELVENGKLDREKAKRIEMLNGIYVLDRAVEIKREYDGVYVLDSDDEBNFLAGFGL 1025
 QY 530 ---NAGYMGWGSARWYCRECAVATWAGRHILR-TAINTARKLGKLVYGGTDSLFT 585
 Db 1026 YANNSYVYGVAKARWYCRECAVATWAGREIYEVKLEKFGKRLVITDGLVAT 1085
 QY 586 Y---DEPKE---NFIKIYKEEG--FEIKLEKYTKLFTFEAKKRYAGLEDGKIDIV 636
 Db 1086 IPGAKPEBEIKKALFEVYVINAKLPGLELELYEGFVNGFFV-TRKYVALIDEGKILIR 1144
 QY 637 GFEAVRGDMCELAKEVQTKVVEIVLKTSEVNAKAVEYRKIVKELEBQKVPDEKVI 696
 Db 1145 GLEIVRRDMSEIAKEVTOARLEILKHGVEBAVRIYKEVTEKLSKYEPPEKVI 1204
 QY 697 SKRLSEYTTAPPHVAAKRMLSGAVRSPGDKIGYIVVKGSGRISORAMPYEMVD 752
 Db 1205 TRPLHEVYKAIQPHVAAKRMLSGAVRSPGDKIGYIVVKGSGRISORAMPYEMVD 752
 QY 753 PSQIDVTVYVDHQIIPALRLIGVFGITEKKLKAASATGQKTLFDFLAKSK 803

Db 1262 KHKYDAEYIENOVLPALVILTEAFYKREKEDLRMQKTKQTGLTAMLNKKK 1312

RESULT 10

C71210

probable DNA-directed DNA polymerase - Pyrococcus horikoshii

C/Species: Pyrococcus horikoshii

C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000

C/Accession: C71210

R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushiida, N.; Oguchi

DNA Res. 5: 55-76, 1998

A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A/Reference number: A71000; MUID:98344137; PMID:9679194

A/Accession: C71210

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-1235 <RAW>

A/Cross-references: GB:AP000007; NID:93236134; PIDN:BAA31074.1; PID:93258391

A/Note: this accession replaces an interim accession for a sequence replaced by GenBank

C/Genetics:

A/Gene: PH1947

C/Superfamily: hypothetical protein PH202

Query Match 22.8%; Score 955.5; DB 2; Length 1235;

Best Local Similarity 23.1%; Pred. No. 5.2e-50;

Matches 299; Conservative 145; Mismatches 291; Indels 559; Gaps 24;

QY 8 VLDSYEVWKEPQVILWIAENGERRVLLIDRSPPRYFALL---APGADPKQVQRIR 63

Db 2 ILDDYITDEDKPIIRIK-KENGEFKVEYDNRPRFYIALLRDSALDEIKITAGRHG 60

QY 64 ALSPPKSIIVEDDKKRYFGRRPRVLRIRTVLPEAVREYELVKNVGDVLEADIRF 123

Db 61 KVVH---IVETEKIQRFKRLPIEVWKLYLEHPQDVPAIRDKIRHRAVVDIFEDYDIF 116

QY 124 AMRLIDHDLPRFTMYRVEAPLENKGFRRDQVLYVSRPEPLYGALAPTKLPDLRL 183

Db 117 AKRVLIDKGL-----TPEMGN-----EKLTF 138

QY 184 AFDEIVSVKSGSPRRDPVIVIAVKTDDGDEVL-----FLAEGKDDKPIREFVEY 235

Db 139 AVDIETLYHEBE-ERKGPFPVIMISYADEGAKVITWKIDLPYEVSVSEEMIKRLIRV 197

QY 236 VKRYDPDIIIVGNNHDPYLLRRARILIGLDVTRFVGAEPITTSVGH---VSPGRL 292

Db 198 IKERDPRVITVYNGDNPDPYLLKRAEKLGIKL-LAGRDSEPKQKMGDSLAIVEIKRI 256

QY 293 NVLDYVAEMPEIKISLEVAEVLGVMKKSERYIIMWEIPDYWDPKPKPLLOQYAR 352

Db 257 HFDLPVIRRTINIPYTYLLEAVEYAI-EGKPEKRYAD--EIAKAMETGEGELERAKSM 313

QY 353 DDVATYVLAELKIPFAIQLSVTGLPRDOYGAMSVGRLEWYLIRAFKMKELPNVVE 412

Db 314 EDATYVYLGEPFMEQALRVQFWDVSRSTGLVEMFLRKAYKERNELAPNCPD 373

QY 413 RPE-----ETVGAIVLEPIRGVHENIAVLDFSSMYNIMIKVNGPDLVRPEKCGE 466

Db 374 EKEVERRLRESYEGGYVEPEKGLWEGIVSLDFRLYPSIITTHVSPDLINRE----- 427

QY 467 CGCME--APVKKFRFCRPGFFVTYLERLLLEAKVRAEMKKYPPSPRYLDEROK 523

Db 428 -GCEBYDVAFVGRHFCDFPGFIPSLIGQLLEEROKIKKMKW-SKDPVEKQLDYQR 485

QY 524 ALKYLVA----- 529

Db 486 AILKILANSILDEMLPIVENKVRVKGDFIDREIENARVYKDGSETILLEVQDKAL 545

QY 530 ----- 529

Db 546 SPNETKSEBLKVVALLIRHRYSGKVGYSIKLSGRRIKITSCHSLFVANGKLVKRGDE 605

QY 530 ----- 529

Db 606 LKPDLLVVBGRLLKLPESQVNLVELLLKLPBEETSNIWMMI PVKGRNPFKMLKTLX 665

QY 530 ----- 529

Db 666 WIRGEGRPRTAGRYLKHLERLGVKLRGCEVLDMESLKRKYLYETLIKRLKTNNS 725

QY 530 ----- 529

Db 726 RAYWVEFNSLRDVSLMIPBELKEMIIGEBRPGKIGFIDVDSFAKLGYISGVEK 785

QY 530 ----- 529

Db 786 DRVFKSKQNVLEDIAKLAELKFGVRRGRGYEVSGKISHAIFRYLAEGRKRIPEFI 845

QY 530 ----- 529

Db 846 SPMDIKAFILKNGNAEELTSTKSELVNLILLNSIGVSDIKIEHEKGYRYRYINK 905

QY 530 ----- 529

Db 906 KESNGDVLDSYEVWKEPQVILWIAENGERRVLLIDRSPPRYFALL---NASYGWSGARW 965

QY 544 YCEBCAAVYANGRHLIRTNARLKLKLVYIGDTSLEFVY---DPEKVE---NFI 595

Db 966 YCEBCASVYANGRQYDLVRRLEARGFVLYIDTGLYATIPGVYDMEVVRALFV 1025

QY 596 KIKKEELG--FEIKLEKVVYRLPFTEAKKRYAGLLEGRIDYGFPAVRDWCLEKEVQ 653

Db 1026 DYNSKLPVLEIEYEFYARGFV-TKKYALIDEGKLVYTGLEIYVRDMSEIATKQ 1084

QY 654 TKVVEILKTSSEVNKAVYRIVKLEBGRVPIEKLVIKWLSKLEEYTTAPHVAA 713

Db 1085 ARVLEAILKHGAVAEAVKIVQDVTEKLTNYEVPPEKLVIEQTRIPNEKAIQPHAVA 1144

QY 714 KRLMSAGYRSPGDKIGYVYVGGGRISORAMPYMWK--DP--SQIDTYVYDHQIIPA 769

Db 1145 KRLMARGIKVYKPGVIGYIYLRDGPISKRA---ISLEEDPFRKHVDAEYIENOVLP 1201

QY 770 ALRILGYGITEKKLKSATGOKTLPDFL-AKKS 802

Db 1202 VERILKAFYKREDLRMQTKQVGLAMIKVKS 1235

QY 1202 VERILKAFYKREDLRMQTKQVGLAMIKVKS 1235

RESULT 11

C69028

DNA-dependent DNA polymerase family B (PolB) - Methanobacterium thermoautotrophicum (s

C/Species: Methanobacterium thermoautotrophicum

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C/Accession: C69028

R/Smith, D.R.; Doucette-Stamm, L.A.; DeJonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func

A/Reference number: A69000; MUID:98037514; PMID:9371463

A/Accession: C69028

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1586 <MTH>

A/Cross-references: GB:AE000888; GB:AE000666; NID:92622304; PIDN:AAB85697.1; PID:9262231

A/Experimental source: strain Delta H

C/Genetics:

A/Gene: MTH1208

Query Match 21.8%; Score 915; DB 2; Length 586;

Best Local Similarity 35.4%; Pred. No. 5.4e-48;

Matches 218; Conservative 110; Mismatches 226; Indels 62; Gaps 12;

QY 1 MTEVVFVLDSSYEVWKEPQVILWIAENG--ERVVLLIDRSPPRYFALLAPGADPKOV 58

Db 1 MEDYRNVLIDIDYTVDEVPVIRLFGKSGNGNEPIAHDRSPREYIAT-----PTDL 54
 Qy 59 AGRIRALSPPKSPITIGVEDDGRKXFGPRVLRIRITVLPEAVREYRELKRVQGVDE 118
 Db 55 DECLRLERLELEKLEVE--MRDLGRPTVIRIERBRHQDPKIRORIRLSVNDIE 112
 Qy 119 ADIRFAMRYLIDHDFP-----FTWVVEAEPLNNMGPRVDKYLKSRPEELXGELA 173
 Db 113 HDIPFYRRYLRIDKSLVPMEELEFGQVVDAPSVTTDRTVTEVYNGRQSGAHG----- 168
 Qy 174 PTLPLRLIADIEVSKOSPPREEDPIYVANK-----TDDGDEVLPIAEGD 224
 Db 169 -----LDILSPDIEVRPHQMPDEKDEIYMLGVAGNMGESVISTAGDHLDEVEVD 222
 Qy 225 DRKPIREPEYVKRYDPIIYGVNNHFMPIYLRRAILGIKLDV-----TRVGA 276
 Db 223 ERRLERFAEIVIDKKPDLVGNDSNDFPPIYTRRALILGELDGLDGSKIRTRRGF 282
 Qy 277 EPTTVHGHVSVGRNLVDLYAEEMPEIKISLEEVAYLGWKSERYIINMEIIPD 336
 Db 283 ANATAIKGTV-----HYDLTPVMRRYMNIDRYTLERYVQELFGSEKIDLPGRLME--- 333
 Qy 337 YMDPPKRPILLQYARDVRYTGLAEKILPFAIQSVYVGLPLDQVGMNSVGRLEWYL 396
 Db 334 YMDRELRLDELFRYSLDDVAVATHIAEKILPLNELTRLVGQPLFDISRMATGQAEWFL 393
 Qy 397 IRAAFPMKELVNRVERPEETFR-----GATVLEPLRGVHENTAVLDFSGMYPNIMIKY 450
 Db 394 VRKAYGELVKNKPSQSDFSRRGRAGVGYKPEKGLHENTIVQDFPSLVSIIISK 453
 Qy 451 NVGPDLTVRGEKCGCGCEMEAPVKHFRCPGPFKTVLERLLERKVRAMKPYR 510
 Db 454 NISPDILTDEES-----ECYAPBYGFRFKSPRGVPSVIEIISERVARKEBKG-SD 508
 Qy 511 DSEYRLDLBEROKALVLANASYGMGSGARWYCECKAKAVTANGHLLIRTAINTARL 570
 Db 509 DPMERKILNVQOELAKRLANMVGYSIFRMYSMECKEALITANGDYIKKITKTAEEF 568
 Qy 571 GLKVIYGDTSLEFVY 586
 Db 569 GFHTVYADTDGFYATY 584

RESULT 12
 T40242
 DNA polymerase delta large chain - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 Accession: T40242
 Borzym, K.; Beck, A.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999
 A:Reference number: T21916
 A:Accession: T40242
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1086 <BOR>
 A:Cross-references: EMBL:AL121815; PTDN:CA58156.1; GSPDB:GND0067; SPDB:SPBC336.04
 A:Experimental source: strain 972h-; cosmid g336
 C:Genetics:
 A:Gene: SPDB:SPBC336.04
 A:Map position: 2
 A:Insertions: 77/1
 C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 18.4%; Score 771.5; DB 2; Length 1086;
 Best Local Similarity 27.3%; Pred. No. 6.7e-39;
 Matches 246; Conservative 160; Mismatches 319; Indels 177; Gaps 30;

Qy 3 EVFETLDSYEVGKEPOVILMGIANGSRVILDRSPRYFALLAPGADPKQYAKI 62
 Db 97 DIFQOIDESEFTSGVSIRLFEVTNGSILVHVAGLPIFYKAPVGRPRMELERT 156
 Qy 63 RAA-----SRKSPITIGVEDDGRKXFGPRVLRIRITVLPE 98

Db 157 QDDATCNGGVLDHCIEEMKENLYFGNGEKSPFKI-----FTNPRILSRANVFR 210
 Qy 99 AVNERELYVNVGVEDVLEADIRFAMRYLIDHDFPFTWVVEAEPLNNMGPRV----- 154
 Db 211 GERNFELFPVGVG-TTFESNTQYLRFMIDDDVGNMGMHILRSKQGRYQNRNSQ 269
 Qy 155 -----DKYLVVSRPEPIYGALAPTLPLRLIADIEVSKOS-SPREERPVY 205
 Db 270 IEAMINVKDLSI-----PAGQW--SMAELKMSDLECGARKGPRPDSIDPVLO 320
 Qy 206 IAVKTDDGEVLPIA-----EGKDRKPIREFVEYVKRYDPIIYV 247
 Db 321 IASIVQYDSDTFPRVNVFCVDTCSQIVGTQVYEFQNAEMLSWSKRVHDPDVLIGY 380
 Qy 248 NNNHFMPIYLRRAILGI-----KLDVTRRVGAEPPTSVHGH-----VSVGRLN 293
 Db 381 NIGCFDIPYLLDRAKSLRINHFPULLGRINHFPVAKETTFSSKAYGTRRESKTSIGRLQ 440
 Qy 294 VDLVYAEEMPEIKISLEEV-AEYLGMKKSERYIINMEIIPYMD--DPKKRPILLO 349
 Db 441 LDMLOVWQDRPFLRSYSINAVCSQFLGQEKED---VHSIITDLONGTADSRRL-LAI 494
 Qy 350 YARDVRYTGLAEKILPFA--IQSVYVGLPLDQVGMNSVGRLEWYLIRAAFPMKELY 407
 Db 495 YCLKDAVLPQRLMDLGMCFVNYTEMARVTCVPPNPLARGOQIVISQLFRKALQHDLYV 554
 Qy 408 PN-RVERPEETFRGATVLEPLRGVHEN-IVLDFSGMYPNIM----- 447
 Db 555 PNIRVNGTDEQEGATVLEPIKGYTDPITLDFSSILPSIQAHNLCYTTLLDSTAE 614
 Qy 448 -----IKNVGP--DLYVRGEKCGCGCEMEAPVKHFRCPGPFKTVLERLLERK 499
 Db 615 LKAKODVDSVTPNDYFVKPHVR-----KGLPITLIDLNLAK 654
 Qy 500 RVRAEMKYPSPSPRYLLDEROKALVLANASYGMGSGARWYCECKAKAVTANGHLL 559
 Db 655 KAKADLKK-ETDPPKAVALDGRQALAKVSNASYGTGATNGNRJPCLAISSVSYGROM 713
 Qy 560 IRTAINIARL-----GLKVIYGDTSLEFVYDPEKVENPKIIEKELG----- 604
 Db 714 IEKTDVVERRYRINGYSHDAVITYGDTSVVVKGVYKLPLEAMKLGEEAANYVSDQFP 773
 Qy 605 ---EIKLEKYRLLPFEAKKRYAGLL-----EDGRIDVGEAVRGMCELAKEYQTV 656
 Db 774 NPIKLEPEKYPF-YLISKSRVAGLPWTRTDYDMOSKIGETVRDNCPLVSVIDTA 832
 Qy 657 VEIVLKTSEVNRNAVYRKIKYLEEGKVPLEKLVIMKLSRLEETTEAPHYVAKEM 716
 Db 833 LRKRLIDQVEGAKLPFTKAVISDLQNKIDMSQVITKALSK--TDVAKKAHVELAEM 890
 Qy 717 --LSAGRVSPGDKIGYIVVG--GGRISQRAV-PYRWKQDSQIDVTYVHQIIPAL 771
 Db 891 RKRDAGAPALGDAVAVIIRGAQGDQFYMRSEDPYVLENNIPIDAKYILENQLSKPIL 950
 Qy 772 RI 773
 Db 951 RI 952

RESULT 13
 S71551
 DNA-directed DNA polymerase (EC 2.7.7.7) KOD, intein containing precursor - Pyrococcus
 N:Contains: DNA endonuclease (EC 3.1.1.-) PI-1; DNA endonuclease (EC 3.1.1.-) PI-11; DN.
 C:Species: Pyrococcus sp.
 A:Variety: strain KOD1
 C:Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 20-Apr-2000
 Accession: S71551
 R:Kakihara, H.; Takagi, M.; Imanaka, T.
 submitted to the EMBL Data Library, March 1994
 A:Description: Cloning and analysis of the DNA polymerase gene from a new hyperthermophilic
 A:Reference number: S71551
 A:Accession: S71551

C:DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - fission yeast (Schizosaccharomyces pombe)
 N:Alternate names: DNA polymerase delta
 C:Species: Schizosaccharomyces pombe
 C:Update: 11-Jan-2000 #sequence_rev15on 11-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T43266
 R:Pigade, G.; Baviar, D.; de Recondo, A.M.; Baladaci, G.
 J. Mol. Biol. 222, 209-218, 1991
 A:Title: Characterization of the POL3 Gene product from Schizosaccharomyces pombe indica
 A:Reference number: S19661; MUID:92012554; PMID:11960723
 A:Accession: T43266
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-1086 <PIG>
 A:Cross-references: EMBL:L07734; NID:g173383; P1DN:AAA35503.1; P1D:g173384
 A:Genetic:
 A:Gene: pol3
 A:Interon: 77/1
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match	18.2%;	Score 763.5;	DB 2;	Length 1086;
Best Local Similarity	27.2%;	Pred. No. 2e-38;		
Matches 245;	Conservative 160;	Mismatches 320;	Indels 177;	Gaps 30;

```

0Y 3 EWFATLDSYSYVWKEPOVIMIGIANENBERVLLIDRSFRFYFALLAPADPKVOARI 62
Db 97 DIFVQOIDSEETEGSVSIRLFCVTDNGSILIVHGFLLFYVKAUVGFRREMLHEFT 156
0Y 63 RAL-----SRKSPYIGVDDRKRYKFGPRRVIRITYPE 98
Db 157 ODDATCNGGVIDHCIIEMKENLYFGQNEKSPFIK-----PTTNPILISANVPER 210
0Y 99 ABBEYELKONVDGVEDVLBEADIRFARVYIHDHLEPPTYRYVAEPLNKQGFV---- 154
Db 211 GERNFELPPVGVY-TTFESNTQYLLRFWIDODVGNMWHILPASKYQPRXONVSNQ 269
0Y 155 -----DKVLYVSREPLVGEALAPTYLDLILAFDEYVSKOG-SPREDDPIYV 205
Db 270 IEAMINYKOLISL-----PABQW---SMAPLHNSFDEICAGRGVPPDSIDPIYQ 320
0Y 206 IA-VKTDDBDEVLFIA-----EKGDDRKPIREFVYVYRVDPIYGY 247
Db 321 IASIVTGYDSTPFVRANVFCVDTCSQIVGTVGYVEFQNGAEMLSMSKFRADVDPYLLGY 380
0Y 248 NNNHPMPVYLRAAILGI-----KLDVTRVGAEEPTSVNH-----VSYPGLN 293
Db 361 NICNFIPIYLDRAKSLRINHFPILGRINHFSVAKETTPSKRYGVRSEKTSIPIRLQ 440
0Y 294 VDLVYDAEMPEIKIKSLSEY-AEYLSYMKCSBRVYIIMWIEPIYWD---DPKXPRLLQ 349
Db 441 LDMLOVQMGDPFLRSYSLNVCSPFLOEKED-----VYSITLDQNGTDSRR--LAI 494
0Y 350 YARDVBARVYLAELIPEA--IOLSYVGLRDOVGAMSVGRFLEVYLLRAAFKCKELV 407
Db 495 YCKLDVILPQRLMDLWCFVYVYTEMARVYVPPNELLARGOIKIVISQLECRALQMDLV 554
0Y 408 PN-RVBERBETRGVIVLEPGRVHEN-IAVLDPSMPTM----- 447
Db 555 PNIRVNGTDEOEGVATIEPIKGYTPYPIYDLSFSLIPMOAHNICYTLTLDNSTVLE 614
0Y 448 -----IKNNP--DILVRGEKCEGCGWENAPVGRPRCPFGFKTVLERLLIARK 499
Db 615 LKLDKQDVSVTPNDYFVKBVH-----KGLPILLDLILNARK 654
0Y 500 RVABEMKPYPPSPSEYLLDERQKALVTYLANSYGMGSGARVYCECKAKAVTAWGRHL 559
Db 655 KAKMDLCK-ETPFFKAVLDGRQALRANSNYSYGTATNGRILPCLATISSYSTSYSGOM 713
0Y 560 IRRIRINIARKL-----GLKVIYGDYDLSLFYVYDPEKVENFIKIKKEELGF----- 604
Db 714 IERTKQDVERRYIENGSHDAVVIYGDYDTSVWVKGFKTLPEAMKIGEEAANVYSDOPF 773
0Y 605 --*EIKLEKRYKVLFFTEAKRRVAGLL-----EDGRIDI VGFEAVRGDMETLAKEVQTV 656

```

```

Db 774 NPKLEFERYFP-XLLISKKRYAGLFWRTRDTYDDKXDSKGIETVRDNCPLVSVVIDTA 832
Qy 657 VEIVLKTSEVNKAVEYRKIVKELEBQKVPLEKVIYKNTLSKLEBYTTBAPHVVAARM 716
Db 833 LRMMIDDPVGAQLFPRKVISDLOKNKIDMSQVITKALSK--TDYAKMAMHVELERM 890
Qy 717 --LSAGYVSPGDKIGYIVIKG--GGRIISOPAM--PYFMVDPQSODIVTYVDHQIIPAAI 771
Db 891 RKRDDGSAAPALDRVAVVIIKGAQGDQFYMSESDPIYVLENNIPIDAKYYLENQSKPLL 950
Qy 772 RI 773
Db 951 RI 952

```

RESULT 15

A41618
 DNA-directed DNA polymerase (EC 2.7.7.7) delta catalytic chain - human
 C/Species: Homo sapiens (man)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #ext_change 10-Sep-1999
 C/Accession: A41618, J35455
 R/Chung, D.W.; Zhang, J.; Tan, C.K.; Davie, E.W.; So, A.G.; Downey, K.M.
 Proc. Natl. Acad. Sci. U.S.A. 88, 11979-11201, 1991
 A/Title: Primary structure of the catalytic subunit of human DNA polymerase delta and epsilon
 A/Reference number: A41618, PMID:92107916; PMID:1722322
 A/Accession: A41618

Query Match	18.2%	Score 760.5;	DB 1;	length 1107;
Best Local Similarity	28.4%	Prod No. 32e38;		
Matches 253;	Conservative 145;	Mismatches 324;	Indels 165;	Gaps 31;

```

Qy 17 GKBPQIITIMGLAENGERRVILIDRSFRPYVYALALPGADPFQVQRR-----ALSLPK----- 69
Db 120 GSVPLVRAFGVLDGCFVCCCHINGFAPYFPYFPAPGCGPHEMMDLQRELINLAISDSRG 179
Qy 70 -----SPIIGVEDEKRR-----YFGR--PRRVLIIRFVLPBEAVREXREIYKQVGVSDV---- 116
Db 180 RELTGAVLAVELCSRESMFGYHGHGSPFLRITVALPRVLVAPARRLLQCGIRVAGLCTP 239
Qy 117 -----LEADIRFAMVYLIDHDLPPTWYRVBEAPLENNKMGFRVQVYL-----YKGRP 164
Db 240 SFAPYEAANDPFIHFVDTDI VGCNWLBPAGKVALRKKEATQCCLEADVLMSDVSHF 299
Qy 165 EPIYGALAPFKTDLKILAFDIEVYSKQG--SPREEDPYIVL-----AVTDDGDSVLEIA 220
Db 300 PEPGMRIRP-----LAVLSFDIECGARKGIRPEPERDPYIQCISGLKMGPEPFLIA 354
Qy 221 -----EGKODRKPPIREFVEYVKRYDPDII VGINNNHDPYLLRR----- 261
Db 355 LTLRCPAILGAKVQSYEKEEDL--IQAWSTRIRIMDEVDITGNIQNDLPYLLSRQOT 412
Qy 262 -----RIIGIKLQDV-----TRRVGAEPITSVHGHVSVGRILNLDLYDAEEMPE 305
Db 413 LKQVQEPFLGVRVAGICSNIRDSFQSKQGRDRDTV-----VSNVGVQNDMLQVLLREYK 468

```

QY 306 IKIKSLBEVA-BYIGVMKSEKSEVIIIMWEIPDYMDPKKRPLLQYARDVR---ATYGL 361
Db 469 ILSHTLNANVSFHFGEOKEDEVQHSI-----ITD-----LQNGNDQTRRLRAYCYL 513
QY 362 AEKILPF-----AIOLSYVTGLPLDQVAMSVMGLEWYLIIRAFKMKELVNR 410
Db 514 KDAVLPILRLERLWLVNAVEMARVGVPLSLSRGQGVKVSQLLRQAMHEGLMPV 573
QY 411 VERPEETRGALVLEPLAGVHE-NIAVLDPSSMYNIMIKYNGEDTLVRPG--EKCEC 467
Db 574 KSEGEDTGTATVLEPLKGYDVPIATLDFSSLYPSIMAHNLCTYLLRPGTAQKLG-- 631
QY 468 GCWEAPEVKHFRRCPPG--FFKT-----VLERLLELRKRVRAEMKKYPPDSPETR 516
Db 632 -----LTEDQFIRTPGTDEFVKTSVRKGLPQILENLSARKAKAKELAK-ETDPLRQ 684
QY 517 LLDERQAKLVLANASVGYMGWSGARWYCRECAVATAMGRHLI-----RTAINIA 567
Db 685 VLDERQALAKVSANSVYGTGAQVKLPCLPSQSVTGFGRQMIKTKQLVESKYTVENG 744
QY 568 RKGLKVIYDPTDSLFTYDPEKVENFIKIKEELGF-----EIKLEKYTKRLFFT 618
Db 745 YSTSAKVYVGTDSVMCRFVSSVAEAMALGREADWVSGHFPSPIRLEFEKYFP-YLL 803
QY 619 EAKRKYAGIL-----EDGRIDIVGFEAVRGDMCELAKEVOTKVEIYVLTSEVNKAVEY 672
Db 804 ISKRRYAGLFFSSRPDAHDRMDCKGLEAVNRDNCPLVANLVTASIRLLIDRDPBGAVAH 863
QY 673 VRKIVKELEBGKVPLEKLVIMKTLISKRLSEYTTTEAPHVAAKRM--LSAGYRVSPGDKIG 730
Db 864 AODVTSIDLCHRIDISQLVITKELTRASDVAGQAHVELABRRMRKRDPGSAPSLGDRVP 923
QY 731 YVIV---KGGRIQORAPYFMWKPDSQIDVTYVVDHQIIPALRI 773
Db 924 YVITSAAGVAAYMKSBDPLFVLEHSLPIDTQYYLEQQLAKPLRI 969

Search completed: Apr 11 23, 2003, 07:29:40
Job time : 61 secs

Delaval, Jan

92138

From: Haddad, Maher
Sent: Wednesday, April 23, 2003 7:17 AM
To: Delaval, Jan
Subject: 10/034,849

Mahe Haddad, 1644
9E12/9D08
306-3472

Jan, .

Please search
1) SEQ ID NO:2 and
2) oligo of seq id no. 2.

Thanks-Mahe

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

5

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:28:09 ; Search time 15 Seconds
(without alignments)
2220.366 Million cell updates/sec

Title: US-10-034-849-2

Perfect score: 4188
Sequence: 1 MTEVFTVLDSYEVGKPE.....LKASATGQKTLFPLAKSK 803

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2434	58.1	784	1	DPO2_AERPE
2	1581	37.8	781	1	DPO2_ARCPU
3	1303.5	31.1	763	1	DPO3_SULSH
4	1287	30.7	764	1	DPO3_SULSO
5	1245.5	29.7	773	1	DPO1_THESQ
6	1236.5	29.5	775	1	DPO1_PYRAB
7	1207.5	28.8	771	1	DPO1_PYRAB
8	1197.5	28.6	775	1	DPO1_THESQ
9	986	23.5	824	1	DPO1_METVO
10	965.5	23.1	1312	1	DPO1_PYRSD
11	955.5	22.8	1235	1	DPO1_PYRHO
12	915	21.8	586	1	DPO1_METH
13	845	20.2	1523	1	DPO1_THERM
14	776	18.5	1671	1	DPO1_PYRKO
15	771.5	18.4	1086	1	DPO1_SCHPO
16	765	18.3	1105	1	DPO1_ORYSA
17	760.5	18.2	1107	1	DPO1_HUMAN
18	757.5	17.9	1106	1	DPO1_BOVIN
19	749.5	17.9	1699	1	DPO1_THESQ
20	749	17.9	1038	1	DPO1_CANAL
21	746.5	17.8	1103	1	DPO1_MESAU
22	743.5	17.8	1105	1	DPO1_MOUSE
23	741.5	17.7	1088	1	DPO1_SOYNA
24	737.5	17.6	1081	1	DPO1_ARATH
25	736.5	17.6	1103	1	DPO1_RAT
26	713	17.0	11702	1	DPO1_THBL1
27	705	16.8	1094	1	DPO1_PLAIF
28	691.5	16.5	1097	1	DPO1_YEAST
29	672	16.0	1668	1	DPO1_METH
30	670	16.0	1634	1	DPO1_METH
31	658.5	15.7	1829	1	DPO1_THESQ
32	655	15.6	1092	1	DPO1_DROME
33	647.5	15.5	872	1	DPO1_SULOH

34	628.5	15.0	875	1	DPO1_SULOH
35	625.5	14.9	959	1	DPO1_AERPE
36	598	14.3	1081	1	DPO1_CABEL
37	594.5	14.2	882	1	DPO1_SULOH
38	579.5	13.8	1462	1	DPO1_HUMAN
39	578	13.8	1465	1	DPO1_MOUSE
40	568	13.6	1451	1	DPO1_RAT
41	539	12.9	3130	1	DPO1_HUMAN
42	526	12.6	3122	1	DPO1_MOUSE
43	524.5	12.5	1015	1	DPO1_EBV
44	522.5	12.5	1339	1	DPO1_TRYBB
45	515	12.3	1009	1	DPO1_HSVSA

ALIGNMENTS

RESULT 1
DPO2_AERPE STANDARD; PRT; 784 AA.
ID DPO2_AERPE
AC 093746;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase II (EC 2.7.7.7).
GN POLB OR APE2098.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
% NCBI_TaxID=5636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RA Ishino Y., Cann I.K.;
RT "Isolation of the genes encoding two alpha-like DNA polymerases from Aeropyrum pernix.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=9310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeyama Y., Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Kosugi H., Hasegawa A., Fukui S., Nagai Y., Nishijima K., Nakazawa H., Takamaya M., Masuda S., Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K., Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
-!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + {DNA}(n).
-!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL, AB017501; BAA7563.1; -;
CC EMBL, AP000063; BAA81109.1; -;
CC HSP, P56689; ITGO.
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR004578; Pol2.
DR Pfam: PF00136; DNA_pol_B; 1.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
DR SMART: SM00486; POLB; 1.
DR SMART: SM00106; DNAPOLB.
DR TIGRFAMs: TIGR00592; pol2; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.

KW Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Complete proteome.
 FT CONFLICT 104 104 E -> K (IN REF. 1).
 FT CONFLICT 306 319 KIGERYTLEMOIG -> R (IN REF. 1).
 FT CONFLICT 355 355 E -> Q (IN REF. 1).
 FT CONFLICT 754 759 LRLOY -> SHTSSN (IN REF. 1).
 FT CONFLICT 763 763 T -> I (IN REF. 1).
 SQ SEQUENCE 784 AA; 89639 MW; F36311C4063B2075 CRC64;

Query Match: 58.1%; Score 2434; DB 1; Length 784;
 Best Local Similarity: 59.1%; Pred. No. 3, 7e-146;
 Matches 469; Conservative 113; Mismatches 189; Indels 22; Gaps 5;

QY 17 GKRPVITWIGINENGERVVLIDRSFRPYFALALPAGADPKQVACRIRALSRPSPIIGVE 76
 DB 3 GSTPVIILMGKADGRVAVVYGFRRPYFVLDPGSGVGLDQALAMIRLSPPSPILSYVE 62
 QY 77 DDKRYFGSPRRVLRIRTVLEAVREYRELKVNVDVEDVLEADIRPAMRYLIDHDLFPF 136
 DB 63 RVRRRRIGREVALKVTLLVPASVREYREAVRRLGCVRVLEADIPFALPFIIDFNLVPM 122
 DB 137 TWYRVAEPLKNGKFRVDKYYLV-----KSRPELYGEALAPLKLJDRLIAFDI 187
 DB 123 RMVVAEVRVAAPHGYSVDRAVYTLSDIREDETRIQEDPLKG-----LRVMAFDI 172
 QY 188 EYVSKQSPRPREDVVIYIAVKTDDGDEVLFIAEGKDDKPIREFVEYVKYDPDIIVGY 247
 DB 173 EYVSKMRTDPKDDPVIMIGQAGSIEILEMDSKCKVIAGFVERVYSIDPDVIVGY 232
 QY 248 NNHFPWPLRLRRARILGIXLDTYRAGAPTSSVHGHSVSGRLANDVAYEEMPEIK 307
 DB 233 NQRFPMPLVBARAVLGVKAVRR-SVEPQGLIGHVSQGRNLNDLDFABELHAKA 291
 QY 308 IKLEEVAVYLGWKKSERVYINMWEIPVWDPKRRPLLQYADVRYATGLAEKIIP 367
 DB 292 VKLEEVADYLGWKKSERVYINMWEIPVWDPKRRPLLQYADVRYATGLAEKIIP 351
 QY 368 FAIQLSYVTGLPDDVGMVSGFLEWYLIRAAFKKELVPRVREPEYTGAIYLEP 427
 DB 352 FGALTSQVSGPLDDVMAVSGFLEWRLIRAAFKKELVPRVREPEYTGAIYLEP 411
 QY 428 RGVENIAVLDFSSMYPNIMIKVNVGPDTLVRPGKCGCGMEAEVKRFRCPGPF 487
 DB 412 PGVHEDIAVLDFASMYPNIMIKVNVGPDTLVRPGKCGCGMEAEVKRFRCPGPF 471
 QY 488 KTVLERILELRKRVBAEMKKYPPDSPEYRLDERQALKVLANASGYMGSGARMYCRE 547
 DB 472 KTVLERILELRKRVBAEMKKYPPDSPEYRLDERQALKVLANASGYMGSGARMYCRE 531
 DB 548 CAAVAVTAMGRHLIRTAIINIAKGLKVIYGDTSLFVYVDPKVENFKIIXELGPEIK 607
 DB 532 CAAVAVTAMGRSIIIRTAIKKAGELDELYIDTSLFVKNDEPKVERLIRVREBELGDIK 591
 QY 608 LEVYVRLPFTAKKRYAGLLDERIDVGEFVAGNEDWELAKESQYVVEYIYLVKSEVN 667
 DB 592 VDVYVRLPFTAKKRYAGLLDERIDVGEFVAGNEDWELAKESQYVVEYIYLVKSEVN 651
 QY 668 KAVEYVRKIVKELEGGKPIKELVYKTSKLEETTEAPHVAAKRM.SAGRVSPD 727
 DB 652 EAVDYVRNITIKLRGGQDMKLVIMKTLIRPSMVEARQPIVTAALLMBAIGKYEPBA 711
 QY 728 KIGYVIVKGGRIQORAMPYFMVDPQIDVYVYVHQIIPALALILIGEGITEKLLKAS 787
 DB 712 KIGYVIVKGGRIQORAMPYFMVDPQIDVYVYVHQIIPALALILIGEGITEKLLKAS 770
 QY 788 ATGQKTLFDFLAK 800
 DB 771 GR-QSTLDFMR 782

RESULT 2
 DPOL_ARCFU STANDARD; PRT; 781 AA.

AC 029753;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN POL OR POLR OR APO457.
 OS Archaeoglobus fulgidus.
 CC Archaea; Buryarchaeota; Archaeoglobi; Archaeoglobales;
 CC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxId=2233;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kervavage A.R., Graham D.E., Kyriides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kitzness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodok A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cooton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.M., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RA "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 350:364-370(1997).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + (DNA)(N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE001070; AAB90741.1; -.
 CC
 CC DR HSSP; P56689; ITGO.
 DR TIGR; AF0497; -.
 DR InterPro; IPR002064; DNA_pol_B.
 DR InterPro; IPR004578; Pol2.
 DR Pfam; PF00136; DNA_pol_B_1.
 DR Pfam; PF01104; DNA_pol_B_exo_1.
 DR PRINTS; PR00106; DNAPOB.
 DR SMART; SM00485; POLBc; 1.
 DR TIGRFAM; TIGR00592; pol2_1.
 DR PROSITE; PS00116; DNA_POLYMERASE_B_1.
 KM Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Complete proteome.
 SQ SEQUENCE 781 AA; 89850 MW; 761C7B18FCC61B8B CRC64;

Query Match: 37.8%; Score 1581; DB 1; Length 781;
 Best Local Similarity: 42.5%; Pred. No. 1, 3e-93;
 Matches 350; Conservative 133; Mismatches 273; Indels 68; Gaps 18;

QY 1 MTEVFTVLDSSEYVNGEPQVITWIGINENGERVVLIDRSFRPYFALALPAGADPKQVQA 60
 DB 1 MERVEGLIDADYETIGGKAVVRLMKDDQG-IFVAYDYNFPDYPYVYI---GVDEIDLNK 56
 QY 61 RIRALSRPSPIIGVEDDKRYGFRPRVLRIRTVLEAVREYRELKVNVDVEDVLEADIRPAM 120
 DB 57 AATSTRREVYIKLSPEKQKTLIGREVGIVYAHHPQVRLKADYLSQG---DVREND 113
 QY 121 IRPAMRYLIDHDLPEFTYRVAEAPLENKNG-----FRVDKYYLVKSRPELYGEALAP 175
 DB 114 IPRVIRLIDKDLACMDGIALIEG-----KGGVYIRSYKIEK-----ESTPRM 157
 QY 176 KLPDLRLIAPDIEVYSKQSGSPRPREDVVIYIAVKTDDGDEVLFIAEGKDDKPIREFVEY 235

Db 158 EPELKMVLDPDCEMLSPFGMEPEKDPPIVIVSKTNDDEIILTG---DEKRIISDFVKL 214
 Qy 236 VKRVDPIIVGYNNHPPMPLRLARRALIGKLVTRRVGAPEPTSYH--GHVSVPRGLN 293
 Db 215 IKSDPDIIVGNDAPFMPFLRKRAERWNLPLD---VGRDGSNVVFRGRPKITGALN 270
 Qy 294 VDLVYAEEMBEIKIKSLAEVAYLVGVAKKSERVLIINMWEIPDYWDPKKKPLLYQVARD 353
 Db 271 VDLVIMRISDIKIKLENAEFLGT--KLEIADIKADLYRWRSREKKK--VLNTARQ 327
 Qy 354 DVRAITYGLAEKILPPAQLSVTGLPLDVGAMSVGFLEVYLIRAAFKMELVPRVER 413
 Db 328 DAITYTLAKELIPMAYELSKIRLPVDDVTRMGKGQVMDLLSEAKKIGELIAPNPEH 387
 Qy 414 PEITYRGAIYLEPLRGVHENAVLDSSMYPNIMIKYVNGVDTLVREPKGEGCCWAP 473
 Db 388 -ASIEGAFVLEPRGHEENACLDPSMYPNIMAFNISPTY----GGRD-DCEAP 440
 Qy 474 EVKRFRCPPGFPEVLERLELRKVRAMKKYPPDSPRYRLDEROKALKYLANASY 533
 Db 441 EVGHKFRKSPDGFPRILRMILKGRRLKVELKNLSPSSRYKLLDITKQTLKLTNSFY 500
 Qy 534 GYMGSGARWYCRCAKAVTAMGRLIRTAIINARKGLKVIYGDTSLPVT--YDPEK 590
 Db 501 GYMGMLARWYCHPCAETITAMGRHIFRTSAKIAESMGFKVLYGDTDSIFVTKAGMTED 560
 Qy 591 VENIKIIEKELEGEIKLEKYYKLFTEAKKRVAGLLEDRIDIVGEAVRGWCELAK 650
 Db 561 VDRILDKHEELPIQIEVDEYSAIFVE-KRYVAGLLEDRILVVGLEVRGWCCELAK 619
 Qy 651 EVQKVEIVLKTSEVNAKAVEYRKIVALEEGKVPLEKVIYMKTLISRLSEYTTAPHV 710
 Db 620 KVGGEVLEVLKKNPEKALSLVQVILRIKEGKVSLEEVYIKGLTKKPKESYMOAHV 679
 Qy 711 VAAKMLSAGRVSPGDKIGYIVYKGGRIQORAMPYFVKD-----PSQ 755
 Db 680 KALKAKREMGIIYVSSKIGYIVYKSGNIGDRAVPIDILIDFGENLRITKSGIETIK 739
 Qy 756 IDVYVYDHOIIPALRIIGYFTEKTKKASANGOKTLPPFLA 799
 Db 740 LDKQYIDNOIIPSVLIRFERGTEASLKSS--QMSLDSFFS 781
 RESULT 3
 DPO3_SULSH STANDARD; PRT; 763 AA.
 AC 005706;
 ID 15-DEC-1998 (Rel. 37, Created)
 AC 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN DNA polymerase III (EC 2.7.7.7) (DNA polymerase B3).
 OS Sulfolobus shibatae.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 NCBI_TaxID=2286;
 RN 11;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97252493; PubMed=9098062;
 RA Edgell D.R., Klenk H.-P., Doolittle W.F.;
 RT "gene duplications in evolution of archaeal family B DNA
 polymerases";
 RT J. Bacteriol. 179:2632-2640(1997).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + (DNA) (N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>)

CC or send an email to license@isb.ch.
 CC -----
 DB EMBL: U02874; AAB53089.1; -
 DB Interpro: IPR002064; DNA_pol_B.
 DB Pfam: PF00136; DNA_pol_B_2.
 DB SMART: SM00486; POLBc_1.
 DB PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.
 KW Transferrase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; 763 AA; 88196 MW; 7A7EFD9A5B9305E2 CRC64;
 SQ SEQUENCE
 Query Match 31.1%; Score 1303.5; DB 1; Length 763;
 Best Local Similarity 35.2%; Pred. No. 7e-76;
 Matches 282; Conservative 164; Mismatches 309; Indels 47; Gaps 14;
 Qy 6 FTVLDSSYEYVKEPOVITWGLAENGERVLLIDRSFRPYFALLAPGADPKOYARIRAL 65
 Db 5 FFLIDFSYEIKDNIPLIYIWSIDEGNSCVVERNFPEYVYEGND--EILENI---59
 Qy 66 SRPKSPIIGVEDDKRYKFGPRVLRIRITVLPEAVREYRELKVNQGVDEVLADIRFAM 125
 Db 60 -RKNCEVLLITTKVKRYKLVNVDALLVQFTPTQIKRCREKIRISIRINGIKSIFPADIRFTM 118
 Qy 126 RYLIDHLPFTWYVYEAELLENMGFRVDKYLIVKSRPEPLVGEALAPKLPDLRIAP 185
 Db 119 RYSIDPDLRFTFPKAEVSEVYLE-GFRACKVYILDKILSHYG-----KIELRAIGI 171
 Qy 186 DIEVYSKQSPREPRDVIVIAVKTDDGDEVLPIAEGKDRKPIREFEVYKRYDDIIV 245
 Db 172 DFOIYSKYGSLNPKRDIIVLISMSKSGSWQFSLDSMDLKIIRKVDYIILNYDDPIIY 221
 Qy 246 GYNNNHDMYLLRRARILGIKLVTRVGAPEPTSYHGVSVPGRLNDLYYAEEMPE 305
 Db 232 VPDVDFHMKYIITERANSLGVKIDIGIKIGSEVSQGTGYHSISGRANDLVGLMNERL 291
 Qy 306 IKIKSLAEVAYLVGVMKKSERYLIINMWEIPDYMDPKRLLQYARDVRAITYGLAEKI 365
 Db 292 TGHIDILEVANYIGISPK-RDSLNTMYEISRYMDDEGNRLVQYSLNKKSIYLLGNFL 349
 Qy 366 LPFAIQLSYVTGLPLDVGAMSVGFLEWYLIRAAFKMELVPRVERPEYR-CAIYVL 424
 Db 350 LSPYSELVKIIGLPLDGLSAVSGNRIEASLIRTAKSEELIPIRMDNPRRSKIKKTVI 409
 Qy 425 EPLRGVHENAVLDSSMYPNIMIKYVNGVDTLVREPKGEGC-GCWEAEVYKRRRCR 483
 Db 410 EPKIGIYSDVYVLDISSVYLSVIRKFNISPTLVK-----GQCDDCYVSTISNYKEKEP 464
 Qy 484 PGFEKTVLERLELRKVRAMKKYPPDSPRYRLDEROKALKYLANASGYMGWSGARV 543
 Db 465 SGLYKTFLELSNIQD-----TRKSVIIEELMSSFYDYIHWINSRW 505
 Qy 544 YCRECAKAVTAMGRHLIRTAIINARKGLKVIYGDTSLPVT-YDPEKVENFKIIEEL 602
 Db 506 YSRILAAVAVDELSEYELGKLVITDILKNSGFEVILLANPLVYVKGSGGKINELLFKINSLV 565
 Qy 603 GFBKLEKVTYKRLPFTFAKRYAGLLEDGRIDIVGEAVRGWCELAKEVQTVYVEIVLK 662
 Db 566 DLMKAKRIYRSLIIL-GNDRVAGLLEGDKIDIRIGKEDRDICEIVRNKRVYVEIILI 624
 Qy 663 TSEYNKAVEYVRKIYKLEEGKVPLEKVIYMKTLISRLSEYTTBAPHVNAKMLLSAGYR 722
 Db 625 SKDYKAVKLVKANVILKLRGGEFDIGELITWHAIEKDFSYDQOLPPVAAARAIOSGYL 684
 Qy 723 VSPGDKIGYIVYKGGRIQORAMPYFVKDPSQIDVTVYVYDHOIIPALRIIGYFTEK 782
 Db 685 ISKOSRIGYIIVGHGVHRAEPFFVKEKNRIDIEYVD-QLRESLKLVLPLGVSEE 743
 Qy 783 KLRASATGQTLVD-FLAKSK 803
 Db 744 SLK--KTNIYDILDMFGASKK 763
 RESULT 4

ID _POS1 SUISO STANDARD; PRT: 764 AA.
 AC P95979; 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase III (EC 2.7.7.7) (DNA polymerase B3).
 GN DPO3 OR SS00081 OR G04041 OR C05_029.
 OS Sulfolobus solfataricus.
 OC Archaea: Crenarchaeota: Thermoprotei: Sulfolobales; Sulfolobaceae;
 CC Sulfobolus.
 CX NCBI_TaxId=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=97055432; PubMed=8899719;
 RA Sengen C.W., Klenk H.-P., Singh R.K., Allard G.,
 RA Liu Q.Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,
 RA Doolittle W.F., Ragan M.A., Charlebois R.L.;
 RA "Organizational characteristics and information content of an
 RA archaeal genome: 156 kb of sequence from Sulfolobus solfataricus
 RA P2.";
 RL Mol. Microbiol. 22:175-191(1996).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332292; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aavay M.J., Chan-Welner C.C.-Y., Clausen I.G., Cutlis B.A.,
 RA De Moers A., Erasuo G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Ndi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.;
 RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 RN [3]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE=97252493; PubMed=908062;
 RA Edgell D.R., Klenk H.-P., Doolittle W.F.;
 RT "Gene duplications in evolution of archaeal family B DNA
 RT polymerases.";
 RL J. Bacteriol. 179:2632-2640(1997).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA] (N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; Y08257; CAA63570.1; -
 CC DR EMBL; AE006648; AAK04041.1; -
 CC DR InterPro; IPR002064; DNA_POL_B.
 CC DR Pfam; PF00136; DNA_POL_B.1.
 CC DR SMART; SMO0466; POLBc.1.
 CC DR PROSITE; PS00116; DNA_POLYMERASE_B; PALS8 NEG.
 CC Transference; DNA-directed DNA polymerase; DNA replication;
 CC KM DNA-binding; Complete proteome
 CC SEQUENCE 764 AA, 88378 MW, 1A4325CTD9A4D672 CRG64;
 OY
 Query Match 30.7%; Score 1287; DB 1; Length 764;
 Best Local Similarity 34.8%; Pseq. No. 7.9e-75;
 Matches 280; Conservative 167; Mismatches 305; Indels 52; Gaps 16;
 DB 6 FTVQSSYEVVKEPPVIWIGIAENGSRVITIDRSFRPYFVALLAGADPKVQVORITRL 65
 DB 5 FFLPDSYVEIKGNPLVYIWSVDDESSVVDNNFPPYIIIV--GNENIIENI--- 59

Accession	Species	Gene	Length (bp)	GC Content (%)	Accession	Species	Gene	Length (bp)	GC Content (%)
Qy 66	SEKSPILLI	CDDEKRCPCRP	187	41.2	Qy 66	SEKSPILLI	CDDEKRCPCRP	187	41.2
Db 60	KONEALOI	TKRYKRIKLNIVDALLO	118	41.2	Db 60	KONEALOI	TKRYKRIKLNIVDALLO	118	41.2
Qy 126	RVLIDHDL	PEPTMYRVEAPLEENKGRV	185	41.2	Qy 126	RVLIDHDL	PEPTMYRVEAPLEENKGRV	185	41.2
Db 119	RISLDPDR	PEPTMYRVEAPLEENKGRV	171	41.2	Db 119	RISLDPDR	PEPTMYRVEAPLEENKGRV	171	41.2
Qy 186	DIETVSKG	SPREBRDPIVIAVKTDD	245	41.2	Qy 186	DIETVSKG	SPREBRDPIVIAVKTDD	245	41.2
Db 172	DPQIYISK	GLSNRPKQPIVMSLSKSG	231	41.2	Db 172	DPQIYISK	GLSNRPKQPIVMSLSKSG	231	41.2
Qy 246	GYNNHDP	WYLLRRARILGILKDVTR	305	41.2	Qy 246	GYNNHDP	WYLLRRARILGILKDVTR	305	41.2
Db 232	YVDSOLP	EMPKYITERASSLGYVIDI	291	41.2	Db 232	YVDSOLP	EMPKYITERASSLGYVIDI	291	41.2
Qy 306	IKIKSLBE	VAYLGWVKSSERVINMWE	365	41.2	Qy 306	IKIKSLBE	VAYLGWVKSSERVINMWE	365	41.2
Db 292	LGHVLDL	IVSNVYLGII--SPSRFS	349	41.2	Db 292	LGHVLDL	IVSNVYLGII--SPSRFS	349	41.2
Qy 366	LEPVALS	YVYVGLPQVQVANSVFE	423	41.2	Qy 366	LEPVALS	YVYVGLPQVQVANSVFE	423	41.2
Db 350	LSTYSEL	VYKIVGLPLDKLSV	409	41.2	Db 350	LSTYSEL	VYKIVGLPLDKLSV	409	41.2
Qy 424	LEPLSGHE	NIAVLDPSSVYPIIMIK	483	41.2	Qy 424	LEPLSGHE	NIAVLDPSSVYPIIMIK	483	41.2
Db 410	IDPKVGI	YTDVYLDISSVYSLVR	465	41.2	Db 410	IDPKVGI	YTDVYLDISSVYSLVR	465	41.2
Qy 484	BPPEFTV	LELRLRKRVALEKKYPP	543	41.2	Qy 484	BPPEFTV	LELRLRKRVALEKKYPP	543	41.2
Db 466	SGLYTPI	DELISNVR-----DS	506	41.2	Db 466	SGLYTPI	DELISNVR-----DS	506	41.2
Qy 544	YCRBCAK	VYVAMGRHLIRALINIA	602	41.2	Qy 544	YCRBCAK	VYVAMGRHLIRALINIA	602	41.2
Db 507	YSRLEIAS	PFDEFSENIIRITIDLI	566	41.2	Db 507	YSRLEIAS	PFDEFSENIIRITIDLI	566	41.2
Qy 603	GREITL	ENYKRLPFTBAKCRYAGL	658	41.2	Qy 603	GREITL	ENYKRLPFTBAKCRYAGL	658	41.2
Db 567	NLDVAVKI	FYKSLILVLD--NNR	621	41.2	Db 567	NLDVAVKI	FYKSLILVLD--NNR	621	41.2
Qy 659	IYLVKTS	EVNAKVEYKIVKELEGKVP	718	41.2	Qy 659	IYLVKTS	EVNAKVEYKIVKELEGKVP	718	41.2
Db 622	ELISKDVK	AKIKVSVYIKLRGEE	681	41.2	Db 622	ELISKDVK	AKIKVSVYIKLRGEE	681	41.2
Qy 719	AGRYVSP	DKIGYIVYKGGRIISOR	778	41.2	Qy 719	AGRYVSP	DKIGYIVYKGGRIISOR	778	41.2
Db 682	SGELISK	OSKSGYIVYKGGRIISOR	740	41.2	Db 682	SGELISK	OSKSGYIVYKGGRIISOR	740	41.2
Qy 779	ITEKTL	--KASATGOQKLPFLAKK	801	41.2	Qy 779	ITEKTL	--KASATGOQKLPFLAKK	801	41.2
Db 741	VNEESL	KKTNTITDILDFGASKK	764	41.2	Db 741	VNEESL	KKTNTITDILDFGASKK	764	41.2

RL Proc. Natl. Acad. Sci. U.S.A. 96:3600-3605 (1999).
 CC - FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
 CC - CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA] (N).
 CC - SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 DR PDB: 1TGO, 22-MAR-99.
 DR InterPro: IPR002064; DNA_pol_B.
 DR InterPro: IPR004578; pol2.
 DR Pfam: PF0136; DNA_pol_B_1.
 DR Pfam: PF0104; DNA_pol_B_exo_1.
 DR PRINTS: PR00106; DNAPOLB.
 DR SMART: SM00486; POLBc; 1.
 DR TIGRFAMs: TIGR00592; pol2.1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B_1.
 DR Transerases; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Hydrolyase; Nuclease; Exonuclease; Multifunctional enzyme;
 KW 3D-structure.
 FT DISULFID 428 442
 FT DISULFID 506 509
 SEQUENCE 773 AA; 89812 MW; F67AF04E875FBE4 CRC64;
 Query Match 29.7%; Score 1245.5; DB 1; Length 773;
 Best Local Similarity 36.1%; Pred. No. 3.6e-72;
 Matches 300; Conservative 148; Mismatches 289; Indels 93; Gaps 21;

QY 695 TSKRLTEYTTEAPHVAARMLASGRVSPGDKGVIVYKGGRTISQRAPEYMKDS 754
 DB 665 QITRDLDKDYATGPHVAARMLARGLKIRPVTIVYKGGSGRGDRIP-FDEFDA 723
 QY 755 --GIDVYVVDH01IPALRIILGFITTEKRLASATGQTLTDFLAKS 802
 DB 724 KHYVDAEYIENQVLPAVERILAPFGYRKEDLRKYOTRQVGLAMLKPKT 773
 RESULT 6
 ID DPOL_PYRFU STANDARD; PRT; 775 AA.
 AG P80051; P95584;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) (Pfu polymerase).
 GN POL OR PF0212.
 OS Pyrococcus furiosus, and
 OS Pyrococcus woesei.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261, 2262;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
 RC SPECIES=P. furiosus; STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RX MEDLINE=93181200; PubMed=8441634;
 RA Umeki T., Ishino Y., Toh H., Asada F., Kato I.;
 RT "Organization and nucleotide sequence of the DNA polymerase gene from
 RT the archaean Pyrococcus furiosus".
 RL Nucleic Acids Res. 21:255-265 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=P. furiosus; STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome".
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 396-418 FROM N.A.
 RC SPECIES=P. furiosus;
 RX MEDLINE=92107689; PubMed=1762925;
 RA Mathur E.J., Adams M.W., Callen W.N., Cline J.M.;
 RT "The DNA polymerase gene from the hyperthermophilic marine
 RT archaebacterium, Pyrococcus furiosus, shows sequence homology with
 RT alpha-like DNA polymerases".
 RL Nucleic Acids Res. 19:6952-6952 (1991).
 RN [4]
 RP SEQUENCE OF 396-502 FROM N.A.
 RC SPECIES=P. furiosus;
 RA Mathur E.J., Adams M.W., Callen W.N., Cline J.M.;
 RL Submitted (NOV-1991) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=P. woesei; STRAIN=DSM 3773;
 RX MEDLINE=98432910; PubMed=9758761;
 RA Dabrowski S., Kur J.;
 RT "Cloning and expression in Escherichia coli of the recombinant his-
 RT tagged DNA polymerases from Pyrococcus furiosus and Pyrococcus
 RT woesei".
 RL Protein Expr. Purif. 14:131-136 (1998).
 RN [6]
 RP SIMILARITY TO OTHER POLYMERASES.
 RX MEDLINE=92253396; PubMed=1579479;
 RA Forterre P.;
 RT "The DNA polymerase from the archaebacterium Pyrococcus furiosus does
 RT not testify for a specific relationship between archaebacteria and
 RT eukaryotes".
 RL Nucleic Acids Res. 20:1811-1811 (1992).
 CC - FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
 CC - CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA] (N).

```

CC -1- SUBUNIT: MONOMER.
CC -1- BIOTECHNOLOGY: Because pfu DNA polymerase exhibits the lowest
CC error rate of any thermostable DNA polymerase studied, it is
CC routinely used for PCR. It is sold by Promega.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: D12963; EMBL: A010147; ALA0336.1;
DR EMBL: U84155; AAB67984.1;
DR FIR: S35543; S35543.
DR HSSP: P56689; ITGO.
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR004578; Pol2.
DR Pfam: PF00136; DNA_pol_B; 1.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
DR PRINTS: PR00106; DNAPOLB.
DR SMART: SM00486; POLBc; 1.
DR TIGRPFAM: TIGR00592; pol2; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
DR Transferrase; DNA-directed DNA polymerase; DNA replication;
DR DNA-binding; Hydrolyase; Endonuclease; Complete proteome.
DR KW
DR SEQUENCE 775 AA; 9013 MW; 5668E3D9F4DD40E CRC64;
SQ
Query Match 29.5%; Score 1236.5; DB 1; Length 775;
Best Local Similarity 36.2%; Fred. No. 1.4e-71;
Matches 301; Conservative 153; Mismatches 283; Indels 95; Gaps 23;
QY 8 VDSSEYVKGKEPOVITWGAENGERVLIIDRSFPFYALLAGADPKOV---AQRIR 63
DB 2 IIDVYITEEGKPYIRLFK-KENGKRIEHDRTFPRYIALLRDSKIEVKKITEERG 60
QY 64 ALSRPSPDIIIGVEDDKKRYGRRPRVLRITVLPRAVEVEYELKVDGYEDVLEADIRF 123
DB 61 KIVR---IVDVEKYEKKFKGKPYTWKLYLHPQDPTIREKREHRAVAVDIFEYDIPF 116
QY 124 AMRYLIDHDLFPFTWYVEAEPLNNKGFVVDKVLVLSREPLYGALAPTLPLRL 183
DB 117 AKRYLIDKGLI-----PMEGE-----BELKITL 138
QY 184 AFDIKYSKSGSPRPERRDPYIVIAVKTDDGDEVL-----FLAEKDKRKRIRREYVY 235
DB 139 AFDIETLVHEGE-EFGSGPITMISYADENBAKVIYTKNIDLPYEVVSSREMIKFLKI 197
QY 236 VKRYPDDIIVGNNHNDMPYLLRRALIGICFDYTRVAGAEPTTSVGH---VSYPGL 292
DB 198 IREKPDIIIVTNGSDPFLPLAKRAEKIGIKLTIGRD-GSEPMQRIQMTAVAEVKGKI 256
QY 293 NVDLVDAEMPEIKISLEEVAVYLGVNKKSESVYINWMEIPDYMDPKRPLLQYAR 352
DB 257 HEDLVHVTIRITINPYTLAVYEA-LFKPKKEKYAA--EIAAMSGSGLERVAKSM 313
QY 353 DQVATYGLAEKILPFAIOLSYVTGLPLDOVGAMSVGRLEMYLIRAFKKEILVNRVE 412
DB 314 EAKATYELGKEPFPMEIOLSRVLGQPLMDVGRSSGTMVEMFLIRAYERNVAVANKKS 373
QY 413 RPE-----ETRGAVILEPARGVHENIAVLDFSSMYNNINIKVNGPDTLVAPGKCEB 466
DB 374 EEEYORLRRESYTGPFVKEPKGLMENIYVLDLPAIYSIIITINVSPTDLNMG----- 428
QY 467 CGGWE-APENVKRRRCRCPGFVTVLELLELRKVRAMKMKYPPSPDEKRLDEKQKL 525
DB 429 CKNYDAPGVGHKCKOIPGFIISLGHLEERKQIKTKMGE-TQDPIELKILDIYQKXI 487
QY 526 KVLAAASYGWMGSGARVYCECAKAVTANGHLIRTA-INIAKKGLKVIYDQDTSLFV 584
DB 526 KVLAAASYGWMGSGARVYCECAKAVTANGHLIRTA-INIAKKGLKVIYDQDTSLFV 584

```

DR PROSITE; PS00116; DNA POLYMERASE B; 1.
 KM Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding; Complete proteome.
 FT VARIANT 532 532 R -> S (IN STRAIN GE23).
 FT VARIANT 553 554 HE -> PN (IN STRAIN GE23).
 FT VARIANT 756 756 K -> R (IN STRAINS GE23 AND ST 855).
 FT CONFLICT 263 263 V -> T (IN REF. 1).
 FT CONFLICT 277 277 A -> T (IN REF. 1).
 FT CONFLICT 281 281 A -> V (IN REF. 1).
 FT CONFLICT 320 320 P -> S (IN REF. 1).
 FT CONFLICT 339 339 Q -> H (IN REF. 1).
 FT CONFLICT 359 359 R -> T (IN REF. 1).
 FT CONFLICT 391 391 K -> N (IN REF. 1).
 SQ SEQUENCE 771 AA; 89496 MW; 110A67045A8A5522 CRC64;
 Query Match 28.8%; Score 1207.5; DB 1; Length 771;
 Best Local Similarity 35.7%; Pred. No. 9.7e-70;
 Matches 292; Conservative 150; Mismatches 285; Indels 91; Gaps 21;
 8 VLDSSEYVVGKPEQVYIWIAGNGERVVLIDRSRPYFALLAPGADPKQV---AQRI 63
 2 ILDADYITENGKPIIRIK-KEKGFKEVDRTPFYIYALKDSDAIDEVKKITAEHNG 60
 64 ALSPPKPIIGVEDDKRYFGRPRVLRITVLEAVREYRELKQVGVEDVLEADIRF 123
 61 KIVR---ITEVEKQKFKLRPIEVMKYLEHPQDVPAIREKIREHNAVVDIPEYDIPF 116
 124 AMRVLIDHDLPFTWYRVEAEPLKNGKFRVDKYLVKSRPEPLVGEALAPTKLPDLRIL 183
 117 AKRVLIDKGL-----TPEMGN-----EELTF 138
 184 AFDIEVYKSGSPRRERDPVIVAVKTDGDEVL-----FLAEGKDDKRPREFEY 235
 139 AVDIETLYHEGE-EGKGPIMISYADEGAKVITWKSIDLPYEAVSSEKEMIKRLVK 197
 236 VKRYDPIIIVGNNHFWPILRRARILGKLDVTRVGAEPPTSVGH---VSPGR 292
 198 IREKOPDVIITYNGDNPFPLTKRAEKLGIKLPGRD-NSEPKQMGQDGLAIVEIKRI 256
 293 NVLDIYAEEMPEIKISLEVAEYLGMKKSERYIIMWEIPDYDDPKRPLLOQYAR 352
 257 HFDFPVIARTINPTYLVAEYVEI-EGKSEKQYAH-EIAEAMERKGLERVAKYSM 313
 353 DDVATYGLAKILPFAITQSVTGLPLDOYGAMSVGRLEEMYLIRAFKKEIVPKNVE 412
 314 EDAKVTPELGEPFMEQLRLVGPWADSRSTGLVEMFLRKAYERNELAPNKP 373
 413 RPE-----ETRGALVLEPLRGVHENIAVLDFSSMYFNIMKYVGPDTLVPRGEKGE 466
 374 EREYERRLRESYBEGYVKEPEKGLMEGLVSLDFRSLYPSIITINVSPTLNR--ENCKE 431
 467 CGCMEAPVKKRFRRCPPGFRTVLERLLELRKRVRAEMKYPDPSPRYLLDEROKALK 526
 432 YDV--APQVGRFCDFPGFIPSLGILNLEROKIKKMKK-SKQPVKSKLLDYRQRAIK 488
 527 VLAAVSYVMGSGARWYCRECAKAVTAMGRHLIRTAIINARKGLKTIYDQTSLEFTY 586
 489 ILAASYGYGYAKARWCKEACASVYAMGQYIDLVRRELSRQFKVLYDTQGLVATI 548
 587 D-----PEKVENFIKIKEELG--FEIKLEKVKRLFTFAKKRYAGLLEDGIDIVG 637
 549 PGAGHEIKAKLKVEYINSLPGLLELVYRGFARGFV-TKKKVALIDEBEGITVIRG 607
 638 FEAVRGDMCELAKYQTVNVEIVLKTSEVNTAVEYVKIYVLEBEGKPIELVIMKTL 697
 608 LEIYARDMSIAKETQAKVLAAILHNGVNDVAVKIVKEVTEKLSKYEIPPEKLVYEQIT 667
 698 KRLSEYTEAPHVVAAKMLSGYVSPGDKIGYIVKGGRISSRAPIYMWK--DSS- 754
 668 RPLSEYKALIGHVAVAKLAKGVAKVKGPMYIGIYVLEGGDGPISGDA--IAIEFPDKK 724
 755 -QIDVTYVVDHQIIPALRLIIGYFGITEKUKASATGQ 791

DB 725 HKYDAEYIENQVLPAYERILRAFGRKEDLKQYKTKQ 762
 RESULT 8
 ID DPOL_THES9 STANDARD; PRT; 775 AA.
 AC 056166;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 OS POL OR POLA.
 OS Thermococcus sp. (strain 90N-7).
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Thermococcus.
 NC NCB1_taxonomy:103799;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96224274; PubMed=8643567;
 RA Southworth M.W., Kong H., Kucera R.B., Ware J., Jannasch H.W.,
 RA Petter F.B.;
 RT Cloning of thermostable DNA polymerases from hyperthermophilic marine
 RT Archaea with emphasis on Thermococcus sp. 90N-7 and mutations
 RT affecting 3'-5' exonuclease activity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:5281-5285 (1996).
 CC -1-CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + (DNA) (N).
 CC -1-SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC CC
 DR EMBL; U47108; AAA8769.1; -.
 DR HSSP; P56689; ITGO.
 DR Interpro; IPR002064; DNA_pol_B.
 DR Interpro; IPR004578; Pol2.
 DR Pfam; PF00136; DNA_pol_B_1.
 DR Pfam; PF03104; DNA_pol_B_exo_1.
 DR PRINTS; PRO0106; DNAPOB.
 DR SMART; SM00486; POLBc; 1.
 DR TIGRFAMs; TIGR00592; pol2; 1.
 DR PROSITE; PS00116; DNA POLYMERASE B; 1.
 DR Transferrase; DNA-directed DNA polymerase; DNA replication;
 KM DNA-binding
 SQ SEQUENCE 775 AA; 89781 MW; 71525D63C0C9CD9D CRC64;
 Query Match 28.6%; Score 1197.5; DB 1; Length 775;
 Best Local Similarity 36.1%; Pred. No. 4.3e-69;
 Matches 301; Conservative 139; Mismatches 296; Indels 97; Gaps 24;
 8 VLDSSEYVVGKPEQVYIWIAGNGERVVLIDRSRPYFALLAPGADPKQV---AQRI 63
 2 ILDADYITENGKPIIRIK-KEKGFKEVDRTPFYIYALKDSDAIDEVKKITAEHNG 60
 64 ALSPPKPIIGVEDDKRYFGRPRVLRITVLEAVREYRELKQVGVEDVLEADIRF 123
 61 TVAVKVR---AEKVQKFKLRPIEVMKYLEHPQDVPAIREKIREHNAVVDIPEYDIPF 116
 124 AMRVLIDHDLPFTWYRVEAEPLKNGKFRVDKYLVKSRPEPLVGEALAPTKLPDLRIL 183
 117 AKRVLIDKGL-----TPEMGN-----EELTF 138
 184 AFDIEVYKSGSPRRERDPVIVAVKTDGDEVLFLAEGKDDKRPREFEY 233
 139 AVDIETLYHEGE-EFGTGPIMISYA--DGESEARVITWKSIDLPYDVVSTEKEMIKRFL 195
 234 EYKRVYDPIIIVGNNHFWPILRRARILGKLDVTRVGAEPPTSVGH---VSPG 290

```

Db 196 RYVREKDPVLITNGDNFPAVYKKRCCELTGKTLGRD-GSEPKIQRMGDRPANEVKG 254
Qy 291 RANVLPDYAEEMPEIKTSLSEVAELTGMKSERVLIINMWEIPDYWDPPKRPILLOY 350
Db 255 RIFEDLYPIRTTINLPYLLEAVEAV-FGKKEKVAE--ETAOAMEGEGLEAVARY 311
Qy 351 ARDDVATYGALEKTLPPALQSLVYTGLPDYGAMSVGFLEWYLIRAAFKMKELVNR 410
Db 312 SMEDAKYTELGEFFPEMAQLSRLIGQSLMDVSSSGNIVEMFLRKAYKKNELAPNK 371
Qy 411 VERPEET-----YRGAVLEBPLRGVHENIAVLDFSSMTNIMIKXNVPDTLVRPGEKCG 465
Db 372 PDERLARRRGYAGGYKBERGLMDNIVYLDPSLIITHNVSPTLVR----- 426
Qy 466 ECGCWE---APEVKHFRRCPPGFVKTLERLELRKVRAMKKYPPDSPEYRLDERQ 522
Db 427 --GCKEYDVAPEVGHKFCDDPPGIPSLGDLLLEBKOKIKRMTA-TVDPLEKKLLDNRQ 483
Qy 523 KALKVLNAYSXGVMGSGARVYCRGCAKAVTAMGRHLRTAI-NIARKLGKXYIGDTPS 581
Db 484 RAIKILANSFYGYGYAKARVYCKECSASTAMGRSEYTEMVIRELEKFGKVLVADTNG 543
Qy 582 LFTY---DPE-----KYENFKIKKELG--PEIKLEKVKYKLPFTAKKRYAGLLEDR 632
Db 544 LHATIPGADAEVYKKAKKEFLKYNPKLPGLBLELGEFYRGFFV-TKKKAVIDESCK 602
Qy 633 IDIVGEAVRGDMCELAKEVOTKVEIYLTSEVNKAVEYRKIVKELEGGVPIEKVY 692
Db 603 ITRGLEIYRDMSEIALETQARVLEAILHGDVEAVARIYKEVTEKLSKYEPKPLVY 662
Qy 693 WKTLSKREBYTTEAPHVAAKRLMSAGRVSPGDKIGVYIKVSGGRISORAMPYFVMD 752
Db 663 HEGITREDRLDYKATGCPHAAVAKRLAAGVIRPGVITGVYIKVSGGRIGRAIPAEF-D 721
Qy 753 PS--QIDVYVYVDHIIIPALRLIGVFGITEKKLKAASATGCKTFDFLAKSK 803
Db 722 FTGHRVDAEYIENQVLEAVERILKAGYRKEDLRYQTKQVGLGAWLKVKGK 774

```

```

RESULT 9
ID DPOL METVO STANDARD; PRT; 824 AA.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2003 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7).

```

```

POL.
Methanococcus voltae.
Archaea; Euryarchaeota; Methanococci; Methanococcales;
Methanococcaceae; Methanococcus.
NCBI_TaxID=2188;

```

```

RN (1)
SEQUENCE FROM N.A.
MEDLINE=95014087; PubMed=7929013;
RA Kohnsky J., Paule S.M., Carinaco M.E., Kaney J.W.;
RT "The DNA polymerase gene from the methanogenic archaeon Methanococcus
RT voltae."
RL J. Bacteriol. 176:6402-6403(1994).

```

```

CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA)(N).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: U33366; AAA72443.1; -
DR HSPB: P56689; ITGO.

```

```

DR InterPro: IPR002064; DNA_POL_B.
DR InterPro: IPR004578; Pol2.
DR Pfam: PF00136; DNA_POL_B.1.
DR PRINTS: PR00106; DNA_POLB.
DR SMART: SM00486; POLB; 1.
DR TIGRFAMs: TIGR00592; pol2; 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B.1.
KW transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding.
SQ
SEQUENCE 824 AA; 96754 MW; 094A630D8C1514FC CRC64;

```

```

Query Match 23.5%; Score 966; DB 1; Length 824;
Best Local Similarity 29.0%; Pred. No. 1.5e-55;
Matches 255; Conservative 159; Mismatches 288; Indels 178; Gaps 25;

```

```

Qy 9 LDSYEVNG-KRPVYIWMGANGSERVYLDSRPYFALLAGADPKVQAGRALS 67
Db 12 IDMYKCKLKEPEINLOKECE-----FRPYF--VDTSEPKIYIYDGLNQ 57
Qy 68 -----PK-----SPITGEDDKKTYFGRR 88
Db 58 EIDKLEBEFENNTSLKVODLTINIEIEKIVSDYILNGKDISVSDFRNK--KERK 114
Qy 89 VLRIPTLPEAVREYRELVRKVDGVEDVLEADIRFAMRYLIDHDLFPFTYRVEAEPLN 148
Db 115 ICKVYKYRNHVKIIRVYKFEFG--KSYEFDPFLRRYIMIDODIVSAKYS-EDNKKIN 170
Qy 149 KMGFRVADKYLVKSRPELVYGEALAPTKLPDLRIIAFDIEVYSKQSGRPREDPVIIV 208
Db 171 -----SIPELNCLIAFMELCYCK-EPNAKDPITIMVL 202
Qy 209 KTDGGEVL-----FIAGCKDRKPIREFVEYVRYPDIIYVGNHHPDMY 256
Db 203 FSKDYQKVIYTKKPFENSEYNGCVYVDEKELIQTIELNQ--DIVIYNGDNDFPY 260
Qy 257 ILRRARILGIKLDYRRRGAPEPT-----SVHGIVSVPRGLAVDLYDAEEMPEIKK 309
Db 261 LKGRANIYELIDPMAANSQOPOIKISKGINKSKIPGIIHIDYIPARKLNLVTR 320
Qy 310 STEEVAEYGVKKSERVIINMWEIPDYWDPPKRPILQVARDVATYGLAEKILPRA 369
Db 321 KLENVQELPKINKE--AVDYGIPRWE--TEDTTLRLAYEDALYTKMKNYFLPLE 375
Qy 370 IQLSVYTGLPDQVAMSVGFLEWYLIRAAFKMKELVNRV-----EPPEYRGALV 423
Db 376 IMFSRIVNPPLDYTSRMNSQWVEFLKRSFQNMISPNRPSSSSYRERAKFSYRGYV 435
Qy 424 LEPLRGVHENIAVLDFSSMYENIMIKXNVPDTLVRPGEKCGCWEAEVGRERRCP 483
Db 436 REPLKGIQEDIVSLDPMSLYPSILISHNISPEVYI-----EKKER-ENNE 480
Qy 484 PGFFKTVLERLELRKVRAMK---KYPDSPEYRLDEROKALKVLNAYSXGVMGS 539
Db 481 LGIIPKTLNELLSRKHKKMLKDKIOKEFDE-EYSRLHEQGSIVKLANSHYGLAP 539
Qy 540 GAWYCRGCAKAVTAMGRHLRTAINIARKLGKXYIGDTPSGLVTDPPKVE----- 592
Db 540 MARMYSDKAEVNTGLKXYIQETIEKAEHFGKVIYADDTGFAFKNDYDKLQKKEEN 599
Qy 593 -----NRIKIEEL--GEPIKLEKVKYR-LFTFAKKYAGL 627
Db 600 DKSDKSNPKLSKELIILTKKFLKGINBELPBGHELEFGRFRGLFTY--KKKYALI 657
Qy 628 LEDGRIDIVGEPAVNGDMCELAKEVOTKVEIYLTSEVNKAVEYRKIVKELEGGVPI 687
Db 658 EDGHLIVVGLVVRDMSNIAKQDOAVIRALLEGVDNIAKKIKITNTIDNKKKNINIK 717
Qy 688 EKLVIKTLSEKLEBYTTEAPHVAAKRLMSAGRVSPGDKIGVYIKVGGGRISORAMPY 747
Db 718 NDLLIHQTQKNIIEYKSTAHIEVAKIKIQGDSGRVGDVYIYIVKSGRSISERA--- 774
Qy 748 FMVNDPSQIDVYVYVDHIIIPALRLIGVFGITEKKLAS 787

```


Db 775 ELLEYAGDYDINYIDNOVLPVIRIMESLGISEDELKNS 814

RESULT 10
 DPOL_PYRSD STANDARD; PRT; 1312 AA.
 AC Q51334; Q51335; Q51336;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) (Deep vent DNA polymerase) [Contains:
 DE Endonuclease PI-PspI (EC 3.1.-.-) (Psp-GDB pol intein)].
 GN POL.
 OS Pyrococcus sp. (strain GB-D).
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_Taxid=69013;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 493-517.
 MEDLINE=94094330; PubMed=8269515;
 Xu M.-Q., Southworth M.W., Mersha F.B., Hornstra L.J., Peiter F.B.:
 "In vitro protein splicing of purified precursor and the
 identification of a branched intermediate.";
 Cell 75:1371-1377 (1993).
 RL C
 CC - FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
 CC - FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
 CC INTERN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY
 CC ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE
 CC INTEIN.
 CC - CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA} (N).
 CC - PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 CC FOLLOWED BY PEPTIDE LIGATION.
 CC - BIOTACHNOLOGY: Used in the PCR method because of its high
 CC thermostability and low error rate. Sold by New England Biolabs.
 CC - SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY
 CC - SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
 CC ENDONUCLEASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U00707; AAA67130.1; -
 CC EMBL: U00707; AAA67131.1; -
 CC EMBL: U00707; AAA67132.1; -
 CC DR HSSP; P56689; ITGO.
 CC DR REBASE; 2619; PI-PspI.
 CC DR InterPro: IPR002064; DNA_pol_B.
 CC DR InterPro: IPR003586; Hedgehog_hintc.
 CC DR InterPro: IPR003587; Hedgehog_hintn.
 CC DR InterPro: IPR002203; intein.
 CC DR InterPro: IPR004042; intein_endonuc.
 CC DR InterPro: IPR004578; Pol2.
 CC DR Pfam; PF00136; DNA_pol_B; 1.
 CC DR Pfam; PF03104; DNA_pol_B_exo; 1.
 CC DR PRINTS; PR00106; DNAPOLB.
 CC DR PRINTS; PR00379; INTEIN.
 CC DR SMART; SM00305; Hintc; 1.
 CC DR SMART; SM00306; Hintn; 1.
 CC DR SMART; SM00486; POLBc; 1.
 CC DR TIGRFAMs; TIGR00592; Pol2; 1.
 CC DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
 CC DR PROSITE; PS50818; INTEIN_C_TER; 1.
 CC DR PROSITE; PS50819; INTEIN_ENDONUCLEASE; 1.
 CC DR PROSITE; PS50817; INTEIN_N_TER; 1.
 CC DR TRANSFERASE; DNA-directed DNA polymerase; DNA replication;
 CC KM

KM DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
 KM Protein splicing; Inton homing.
 FT CHAIN 1 492 DNA POLYMERASE, 1ST PART.
 FT CHAIN 493 1029 ENDONUCLEASE PI-PSP1.
 FT CHAIN 1030 1312 DNA POLYMERASE, 2ND PART.
 SQ SEQUENCE 1312 AA; 152852 MW; B6251805641D26A CRC64;
 Query Match 23.1%; Score 965.5; DB 1; Length 1312;
 Best Local Similarity 21.9%; Pred. No. 5; 8e-54;
 Matches 300; Conservative 154; Mismatches 282; Indels 635; Gaps 23;
 Db 8 VLDSYVYVKEPQVITWIGIAENGERYVLLDRSPFPYFVLLAPGADPKQV-----AQRIR 63
 Db 2 ILDAADYITDEGKPIRIKFK-KENGSEFVVEYDRNRPYITALLKDDSDIDVRKITLIERHG 60
 Db 64 ALSRPSPIIGVEDDKRKRYGPRPRVLRITVLRPEAVREYRELKVNVDGEVLEADIRP 123
 Db 61 KIVR-----IIDAERKVRKKFGLRPLEVWRLYFEHPDVPALRDKIREHSAVIDIFEVDI 116
 Db 124 AMRYLIDHDLFPPTWRVVEAEPLKMKGFVVDVYLVKSRPELYGBALPTKLPDLRL 183
 Db 117 AKRYLIDKGLI-----PMEGD-----EBIKL 138
 Db 184 AFDIEVYSKQSPRPERRPVIVIAVKTDGDEVL-----PIAGKDRKDPREFVEY 235
 Db 139 AFDIEITLHGE--EFAKGPITIMSYADEEAKVITMKDLPVEVYSSREMIKRFKV 197
 Db 236 VKRVDPIIVGYNHNEHDEWYLLRRALIGIKLDVTRRGAEPPTSVGH---VSYGRL 292
 Db 198 IREKDPVILITNGDSFDLYLVKRAKIGIKLPLGND--GSEPRKQRLGDMTAVEIKRI 266
 Db 293 NDLYDAEEMPEIKSLSEVAEYLGWKKSRVLIIMWEIPYMDPKRPLLIQYAR 352
 Db 257 HFDLYHIVIRRTIMLPYTLAEVYEA--FGKPKKVVAN--EIAEAVETGKGLERVAKYSM 313
 Db 353 DDVATYGLAKELIPALIOISYVTGLPLDQVAMSGFRELWLLIAAFMKELPRVVE 412
 Db 314 EDKATYELREFFEPMEAOQLSRVGPDPVSSSTGNLVEWLLKRAVKNELAPKPD 373
 Db 413 RPE-----ETRGATVLEPLRGVHENIAVLDPSSWYPMIKYVNPDTLVPRGEKGE 466
 Db 374 EREYERLRISYAGVYKEBEKGLMEGLVSLDFRSLPSIIITHNVSPDLNNE----- 427
 Db 467 CGGCE--AEVYKRRPRCPGPFKTVLERLLKRVRAEMKPPDPSEYLLDEROK 523
 Db 428 -GCREYDVAPEVGHKFCQDPGFIPSLIKRLDROEIKKMKKA--SKDPIEKMLDYROR 485
 Db 524 ALKYLIA----- 529
 Db 486 AIKILANSILPEEWPLIKNGKVIKIRIGDFVDGLMKANOGKVKKTGTEVEVAGIHAF 545
 Db 530 ----- 529
 Db 546 SPDRKSKARVMAKVAIVIRHRSYGNVRIYVANGRKITITEGHSFLVYRNGDLVEATGED 605
 Db 530 ----- 529
 Db 606 VKIGDLLAVRSVNLPEKERLNIIVELLNLSPEDIEDILTI PVKGRKNFPKGMRLTLR 665
 Db 530 ----- 529
 Db 666 WITGEKRVATASRYLRLHLENLGIIRLKIGYDIIDKEGLEKRYLYEKLVDVRYNGNK 725
 Db 530 ----- 529
 Db 726 RYLVEFNAVADVIALMPEBELKEMRIGTNNGFRMGTFTVIDDPAKLGLGYVBSGARK 785
 Db 530 ----- 529
 Db 786 WKQGTGMSYTVLVYNENDEVLDMEHLAKKFPKVKRGKGVVEIPKMAVITFESLGGT 845
 Db 530 ----- 529

```

Db 846 LAENKAVPEVITFTSSKGVNMAFLSGYFIGDGVDPKSRVLTSTSELLVGLVLLNSIG 905
Qy 530 ----- 529
Db 906 VSAIKGDSGVYRVYVNEELKFTYRKKKNVSHIVPKDILKETFGKVFQKNISYKPF 965
Qy 530 ----- 529
Db 966 RELVNGKLDREKAKRIEMLNGDIVLDRVVEIKREYDGYVYDLSVDEENFLAGEFEL 1025
Qy 530 ---NAGYMGWGSARWYCRECAVATAMGRHLIR-TAINIRKLGKLVITGDDISLFVT 585
Db 1026 YAHNSYGYGYAKARWYCKEACASVTAMGREYIEFVRKELBEKGFVLYIDTDGLVAT 1085
Qy 586 Y---DEKVE---NFIKIKKEELG--FEIKLEKYKRLFFTEAKKRYAGLLBGRIDIV 636
Db 1086 IPGAKPEEIKKKALFEVDYINAKLPGELLEVEYEGFYVRGFFV-TKKYALIDEGKTIIR 1144
Qy 637 GFEAVRGDWCELAKXEVOKVWEIYLTKESEVNAKVEYKIVALEEGKVPTEKLYWKTL 696
Db 1145 GLEIVKRDMSIAKETQAKVLEALKGNVEANVKLVKSEVTEKSKSEIPEKCLVIEQI 1204
Qy 697 SKLEETTEAPHYVAAKMLSAQYVSPQDKIGYIVYKGGRIISQRAWPFYMYKD--- 752
Db 1205 TRPLHEKAIQPHVAVARLARVAKYRPFQWVIGYIVLKGDPISKRA---TLAEPDRLR 1261
Qy 753 PSQIDVYVYVHQIIPALRIILGYFGITEKELKASATQCKTLFDELAKSKR 803
Db 1262 KHKYDAVYIENQVLPVAVIRILEAFGRKEDLWMQTKQTKGLTAMLNKKK 1312

```

```

RESULT 11
ID POL PYRHO STANDARD; PRT; 1235 AA.
AC OS9610;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) (Contains: Pfu pol intein (Pfu Pol I
GN POL OR PH1947 OR PH87047.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CX Pyrococcus.
NM NCBI_TaxID=53953;
RP SEQUENCE FROM N.A.

```

```

MEDLINE=98344137; PubMed=9679194;
Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hogooyama A., Nagai Y.,
Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kushiida N., Oguchi A.,
Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
Maeuchi Y., Shizuya H., Kikuchi H.;
"Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3";
RNA Res. 5:55-76(1998).

```

```

RT DNA Res. 5:55-76(1998).
CC -I- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N)
CC -I- PFM. THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -I- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL, AP000007, BAA31074.1; -.
DR HSSP, P56689, 1TGO.
DR InterPro: IPR002064, DNA_pol_B.
DR InterPro: IPR003586, Hedgehog_hinc.
DR InterPro: IPR003587, Hedgehog_hinc.
DR InterPro: IPR002203, Intein.
DR InterPro: IPR004042, Intein_endonuc.
DR InterPro: IPR004578, Pol2.
DR Pfam: PF00136, DNA_pol_B_exo_1.
DR Pfam: PF03104, DNA_pol_B_exo_1.
DR PRINTS: PR00106, DNAPOLB.
DR SMART: SM00305, Htnc; 1.
DR SMART: SM00306, Htnc; 1.
DR SMART: SM00486, PolBc; 1.
DR TIGRFAMs: TIGR00592, pol2; 1.
DR PROSITE, PS00116, DNA_POLYMERASE_B_1.
DR PROSITE, PS00818, INTEIN_CTER; 1.
DR PROSITE, PS00819, INTEIN_ENDONUCLEASE; 1.
DR PROSITE, PS00817, INTEIN_NTER; 1.
DR TRANSFERASE; DNA-directed DNA polymerase; DNA replication;
DR DNA-binding; Autocatalytic cleavage; Protein splicing;
DR Complete proteome.
DR CHAIN 1 492 DNA POLYMERASE, 1ST PART (POTENTIAL).
DR FT CHAIN 493 952 PHO POL INTEIN (POTENTIAL).
DR FT CHAIN 953 1235 DNA POLYMERASE, 2ND PART (POTENTIAL).
DR SQ SEQUENCE 1235 AA; 143086 MM; 73CCTAA14873CE4 CRC64;

```

```

Query Match 22.8%; Score 955.5; DB 1; Length 1235;
Best Local Similarity 23.1%; Pred. No. 2.3e-53;
Matches 299; Conservative 145; Mismatches 291; Indels 559; Gaps 24;

```

```

Qy 8 VLDSSEYVNGKEPVIITWIGIAENGERVYLIDRSPRYFAYL---APGADPQVORUR 63
Db 2 ILADADYTDGKPIIRIK-KENGEFVNEYDRFRYIYALLRDSALIDEIKITQNRG 60
Qy 64 ALSRPSPIITGVEDDKKYPFRPRVLRITVLPENAREYELVANNVDYEDVLENDIF 123
Db 61 KVA---IVETEKQKRFLEKPLVWKLYLHPQVPAIRKREHPAVADVFEYDIF 116
Qy 124 AMRYLIDHDLFPPTVRYVAEAPLENKMGFRVDKVLVSRBPPLYGBALPTKLPDLRL 183
Db 117 AKRYLIDKGL-----TPMEGN-----EKLTEL 138
Qy 184 AFDIEVYSKQSPRPREDPVIVAVKTDGDEVLT-----FLAEGDKRKEFVEY 235
Db 139 AVDIETLYHEGE-EFGKGPVIMISYADSEGAKVITWKIDLPYVAVSSREMIKRLIRV 197
Qy 236 VKRYDPDIIVYNNHNPMPYLLRRARILGIKLVDYRRVGAEPPTSVYGH---VSPGRRL 292
Db 198 IKEDQDPVITITNGNDFPFLRLRAEKIGITL-LIGRDNSEPMQKMGDSLAVEIKGRI 256
Qy 293 NVLDYDAEMBEIKSLSEVAEYGVMMKSRVYINMMELPDVMDPKKRULLQYAR 352
Db 257 HFDLPYIKRTINLTPLYLLENVYAL-FGKPKENYAD--EIAAKMETSGLERVAKYEM 313
Qy 353 DVRAATYGAELKILPVALQSLVYTGFLPDQYGMASGFLLEYLILRAAFKKELVNRVY 412
Db 314 EDKAVTYTEGREFPEBAQLARLVGPVMDVSSSGNLVPEFLRKAYEENELAPKED 373
Qy 413 RPE-----ETRYGALVLEPLRGVHENIAVLDFSSMYPNIMKYVGPDTLVBPGEKCE 466
Db 374 EKEVERLRLESYEGYVPEPEKGLMEGIVSLDPRSLYPSIITITNVSPDTLNE----- 427
Qy 467 CGCWE---APEVGHFRRCPPGFFKTYLLETLERKVAEMKKYPPDSPSEYVLLDEROK 523
Db 428 -GCSEYDVAVPKGFHFCDFPGLPILGLLEBRQKIKKWK-E-SKDPVEKLLDYRQR 485
Qy 524 ALKTLA----- 529
Db 486 AKTLANSILPDEWLPVVENKFAVFKIGFDIREIEENARVYKRDGTEILVKKLAL 545
Qy 530 ----- 529

```

Db 546 SPNETKSELKVKVALIRHRYSGKVSILKSGRRITKTSGHSLFSGNKLKVRGDE 605
 QY 530 ----- 529
 Db 606 LKPDLLVVPGRRLKPSKQVLANIVELLKLPBEETSNIYMMIPVKGRKPFKMLTLY 665
 QY 530 ----- 529
 Db 666 WIFEGEPRTAGRYLKHLELGYVVKLRGCEVLWDESLKRYRKYETLLKNLKNNGNS 725
 QY 530 ----- 529
 Db 726 RAYWERNSLRDVVSILMPIELKEMIIGEPGPKIGFIDVDSEFAKLLGYIISGDEVK 785
 QY 530 ----- 529
 Db 786 DRVFKSKQNVLEDAKLAELKFGVKRGRGYIEVSGKISHALFRVLAEGKRIPFT 845
 QY 530 ----- 529
 Db 846 SPMDIKVAFKLGNGNNEELFTSTKSELVNOVLILNLSIGVSDIKIEHEKGYRVYINK 905
 QY 530 ----- 543
 Db 906 KESSNGDIVLDSVESIEVEKEYGYVDLSDENENFLVGFGLYAHNSYGYGAKARW 965
 QY 544 YCRCAKAVTAMGRHLIRTAIINARKGLKITYGDTDSLFTTY----DPEKVE---NFI 595
 Db 966 YCKEASASVTAMGRQYIDLVRRELBARGPKVALYDGLVYTTIGVXDMWEVKRRALEFV 1025
 QY 596 KIIKEELG--FEIKLEKYYKLFTEAKRYAGLLEDRIDIVGFEAVRGDMCELAKEVQ 653
 Db 1026 DYNISKLPVLELEBEYGYANGFV--TKKKYALIDBEKITYRGLLEYVRMSIAKETQ 1084
 QY 654 TKVVEILKTSVANKAVEYKVIKELEBKVPIEKVIWMTLSKRLSEYTTAPHHVAA 713
 Db 1085 ARVLEALIKHGNVEAVKIVADVTEKLTNYVEPEKLIYEQITRPINEYKALIPHHVAA 1144
 QY 714 KRMLSAGRVSPGDKIGVIVKGGRIISORAMPYFMVK--DP--SQIDVTYVVOHIIIPA 769
 Db 1145 KRLMARGIKVAPGNVIGIVIRGDSPIKSKRA--ISIEFPBRKHGDAEYIENQVLP 1201
 QY 770 ALRILGYFITEKLLKASATQKTLFDFL-AKKS 802
 Db 1202 VERILKAFGYREDLNRKQTKQVGLGAMIKYKKS 1255

SUBT 12
 L-METTH STANDARD; PRT; 7586 AA.
 AC 027276;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN POL OR POLB1 OR MTH1208.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=8037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
 RA Aldridge T., Baehrzaadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pochler B., Qiu D.,
 RA Spadator R., Viscare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Carnuso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA Doughtall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics.";

RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC +[DNA](N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE000889; AAB85697.1; -
 DR HSSP: P56689; ITGO.
 DR InterPro: IPR002064; DNA_pol_B.
 DR Pfam: PF00136; DNA_pol_B_1.
 DR Pfam: PF03104; DNA_pol_B_exo; 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR SMART: SM00486; POLB; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 KW Transferrase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Complete proteome.
 SQ SEQUENCE 586 AA; 67965 MW; D406B5BC399B951A CRC64;

Query Match 21.8%; Score 915; DB 1; Length 586;
 Best Local Similarity 35.4%; Pred. No. 3, 4e-51;
 Matches 218; Conservative 110; Mismatches 226; Indels 62; Gaps 12;

QY 1 MTEVFTVLDSSYEWGKEPVIIMGIAENG--ERVVLIDRSRPFFVALLARGADPKOV 58
 Db 1 MEYRWVLDDIDVYTVDEVPVIRLFGKDKSGGNEPIIADRSRPFTIYAL-----PTDL 54
 QY 59 AGRIRALSRKPSPIIGVEDDKRYFGRRPVLRIRTVLPEAVREYELVKNVGDVLE 118
 Db 55 DECLRELELELEKLEVKR--MEDLRPFTEVIRIEFRHPQDVKIRIDRLDESVDIRE 112
 QY 119 ADIRFARMYLIHDLRP-----FTWYVRAEPLLENKMGKRVKIVYKSRPELYGEALA 173
 Db 113 HDIPFRARYLIDKSIIVPMEELEFGVVEVDSAPSVTITDVTVEVGRVOSTGSGAHG---- 168
 QY 174 PTKLPDLRIAPDIEVYVSKSGSPRPDPVIVYAVK-----TDDGDEVLFIAEGKD 224
 Db 169 -----LDLISFPIEVRANPFGMPDPEKDELIMIGVAGNMGYESVISTAQDHLDFVEVED 222
 QY 225 DRKPIREFEVYKRYDPIIIVGYNHNFDMVYLLRRARILGILDV-----TRVGA 276
 Db 223 ERELLERFAEIVIDKKPDILVGNSDNDFPYITRRAAILGAEILDGMOSKIRTVRRGF 282
 QY 277 EPTTSVHGHSVGRANLVLDYDAEMPELKISLEBAVEYLGVMKKSERVIIINMEIPD 336
 Db 283 ANATAIKGV-----HVDLYPVMRRYMILDRYTLERVYQELGEBEKIDLPGRDLWE--- 333
 QY 337 YMDPCKRPLLLQYARDVATYGLAEKILPFAIOLSVYTLGPLDVGAMSVGFRLEWYL 396
 Db 334 YMDDELRLDELFRYSDDVAVATIRIAEKILPLNLELTRLVGQPLFPISRMATQOQEWFL 393
 QY 397 IRAAFKKEKLVNRRVERPEETVR-----GAVLEPLRGVHENIIVLDFSSMYPNIMIX 450
 Db 394 VRKAYQYGLVLPKPSQSDPSRGRRAVGCVYKEPEKGLHENVQFPDFSLVPSIILSK 453
 QY 451 NVGPDILVRPGEKCGECGEAPEVHKRRRCPPGPFKYTLERLRLKRVRAEMKKYPP 510
 Db 454 NISPDILTDEES-----ECVVAPEYGRFRFKSPRGVPSVIGILBERVRIKEMKG--SD 508
 QY 511 DSEVEYRLDEROKALKVLANASYGVMGSGARVYCEBCKAVTAMGRHLIRTAIINARKL 570
 Db 509 DPEKRIKLVNQALKRLANTMTGVYGYSRFRYSCECAAITAMGRDITIKTITKTAEEF 568
 QY 571 GLKVIYGDTSLSFTY 586
 Db 569 GFHTVYADTGGFYATY 584

RESULT 13
ID POLYMERASE STANDARD; PRT; 1523 AA.
AC P74918;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) [Pol Tfu] [Contains: Endonuclease PI-Tfu
DE (EC 3.1.-.-) (Tfu pol-1 intein); Endonuclease PI-TfuII (EC 3.1.-.-)
DE (Tfu pol-2 intein)].
GN POL.
OS Thermococcus fumicolans.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
NCBI_TaxId=46540;
NM 11
NP SEQUENCE FROM N.A.
Camion M., Querellou J.;
Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP CHARACTERIZATION OF INTEINS.
RC STRAIN=ST557;
RX MEDLINE=20112788; PubMed=10644683;
RA Saveri I., Ozanne V., Dietrich J., Masson J.-M.;
RT "Inteins of Thermococcus fumicolans DNA polymerase are endonucleases
RT with distinct enzymatic behaviors.";
RL J. Biol. Chem. 275:2335-2341 (2000).
CC -1- FUNCTION: PI-TfuI recognizes and cleaves a minimal sequence of 16
CC base pairs (bp) on supercoiled DNA with either Mn(2+) or Mg(2+) as
CC cofactor. It cleaves linear DNA only with Mn(2+) and requires a
CC 19-bp minimal recognition sequence. The optimal temperature for
CC activity is 70 degrees Celsius.
CC -1- FUNCTION: PI-TfuII is a highly active homing endonuclease using
CC Mg(2+) as cofactor. Its minimal recognition and cleavage site is
CC 21 bp long either on linear or circular DNA substrates. Its
CC endonuclease activity is strongly inhibited by the 3' digestion
CC product, which remains bound to the enzyme after the cleavage
CC reaction. The optimal temperature for activity is 70 degrees
CC Celsius.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N).
CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE TWO INTERVENING REGION
CC (INTEINS) FOLLOWED BY PEPTIDE LIGATION
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -1- SIMILARITY: IN THE INTEIN SECTION, BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z69882; CA93738.1; -
DR HSSP; P56689; 1TCO.
DR REBASE; 4500; PI-TfuI.
DR InterPro; IPR002064; DNA_pol_B.
DR InterPro; IPR003586; HcdGehog_hintc.
DR InterPro; IPR003587; HcdGehog_hintn.
DR InterPro; IPR002203; intein.
DR InterPro; IPR004578; intein_endonuc.
DR InterPro; IPR004578; pol2.
DR Pfam; PF00136; DNA_pol_B; 2.
DR Pfam; PF01044; DNA_pol_B_exo; 1.
DR PRINTS; PRO0379; INTEIN.
DR SMART; SM00305; HINC; 2.
DR SMART; SM00306; HINC; 2.

DR SMART; SM00486; POLB; 1.
DR TIGRFAMs; TIGR00592; pol2; 1.
DR PROSITE; PS00116; DNA POLYMERASE B; FALSE_NEG.
DR PROSITE; PS50818; INTEIN C TER; 2.
DR PROSITE; PS50819; INTEIN_ENDONUCLEASE; 2.
DR PROSITE; PS50817; INTEIN_N TER; 2.
KW transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Hydrolyase; Endonuclease; Autocatalytic cleavage;
KW Protein splicing; Intron homing; Magnesium; Manganese;
FT CHAIN 1 406
FT CHAIN 2 766
FT CHAIN 3 900
FT CHAIN 4 901
FT CHAIN 5 1282
FT CHAIN 6 1283
FT CHAIN 7 1523
FT CHAIN 8 1523
FT CHAIN 9 1523
FT CHAIN 10 1523
FT CHAIN 11 1523
FT CHAIN 12 1523
FT CHAIN 13 1523
FT CHAIN 14 1523
FT CHAIN 15 1523
FT CHAIN 16 1523
FT CHAIN 17 1523
FT CHAIN 18 1523
FT CHAIN 19 1523
FT CHAIN 20 1523
FT CHAIN 21 1523
FT CHAIN 22 1523
FT CHAIN 23 1523
FT CHAIN 24 1523
FT CHAIN 25 1523
FT CHAIN 26 1523
FT CHAIN 27 1523
FT CHAIN 28 1523
FT CHAIN 29 1523
FT CHAIN 30 1523
FT CHAIN 31 1523
FT CHAIN 32 1523
FT CHAIN 33 1523
FT CHAIN 34 1523
FT CHAIN 35 1523
FT CHAIN 36 1523
FT CHAIN 37 1523
FT CHAIN 38 1523
FT CHAIN 39 1523
FT CHAIN 40 1523
FT CHAIN 41 1523
FT CHAIN 42 1523
FT CHAIN 43 1523
FT CHAIN 44 1523
FT CHAIN 45 1523
FT CHAIN 46 1523
FT CHAIN 47 1523
FT CHAIN 48 1523
FT CHAIN 49 1523
FT CHAIN 50 1523
FT CHAIN 51 1523
FT CHAIN 52 1523
FT CHAIN 53 1523
FT CHAIN 54 1523
FT CHAIN 55 1523
FT CHAIN 56 1523
FT CHAIN 57 1523
FT CHAIN 58 1523
FT CHAIN 59 1523
FT CHAIN 60 1523
FT CHAIN 61 1523
FT CHAIN 62 1523
FT CHAIN 63 1523
FT CHAIN 64 1523
FT CHAIN 65 1523
FT CHAIN 66 1523
FT CHAIN 67 1523
FT CHAIN 68 1523
FT CHAIN 69 1523
FT CHAIN 70 1523
FT CHAIN 71 1523
FT CHAIN 72 1523
FT CHAIN 73 1523
FT CHAIN 74 1523
FT CHAIN 75 1523
FT CHAIN 76 1523
FT CHAIN 77 1523
FT CHAIN 78 1523
FT CHAIN 79 1523
FT CHAIN 80 1523
FT CHAIN 81 1523
FT CHAIN 82 1523
FT CHAIN 83 1523
FT CHAIN 84 1523
FT CHAIN 85 1523
FT CHAIN 86 1523
FT CHAIN 87 1523
FT CHAIN 88 1523
FT CHAIN 89 1523
FT CHAIN 90 1523
FT CHAIN 91 1523
FT CHAIN 92 1523
FT CHAIN 93 1523
FT CHAIN 94 1523
FT CHAIN 95 1523
FT CHAIN 96 1523
FT CHAIN 97 1523
FT CHAIN 98 1523
FT CHAIN 99 1523
FT CHAIN 100 1523
FT CHAIN 101 1523
FT CHAIN 102 1523
FT CHAIN 103 1523
FT CHAIN 104 1523
FT CHAIN 105 1523
FT CHAIN 106 1523
FT CHAIN 107 1523
FT CHAIN 108 1523
FT CHAIN 109 1523
FT CHAIN 110 1523
FT CHAIN 111 1523
FT CHAIN 112 1523
FT CHAIN 113 1523
FT CHAIN 114 1523
FT CHAIN 115 1523
FT CHAIN 116 1523
FT CHAIN 117 1523
FT CHAIN 118 1523
FT CHAIN 119 1523
FT CHAIN 120 1523
FT CHAIN 121 1523
FT CHAIN 122 1523
FT CHAIN 123 1523
FT CHAIN 124 1523
FT CHAIN 125 1523
FT CHAIN 126 1523
FT CHAIN 127 1523
FT CHAIN 128 1523
FT CHAIN 129 1523
FT CHAIN 130 1523
FT CHAIN 131 1523
FT CHAIN 132 1523
FT CHAIN 133 1523
FT CHAIN 134 1523
FT CHAIN 135 1523
FT CHAIN 136 1523
FT CHAIN 137 1523
FT CHAIN 138 1523
FT CHAIN 139 1523
FT CHAIN 140 1523
FT CHAIN 141 1523
FT CHAIN 142 1523
FT CHAIN 143 1523
FT CHAIN 144 1523
FT CHAIN 145 1523
FT CHAIN 146 1523
FT CHAIN 147 1523
FT CHAIN 148 1523
FT CHAIN 149 1523
FT CHAIN 150 1523
FT CHAIN 151 1523
FT CHAIN 152 1523
FT CHAIN 153 1523
FT CHAIN 154 1523
FT CHAIN 155 1523
FT CHAIN 156 1523
FT CHAIN 157 1523
FT CHAIN 158 1523
FT CHAIN 159 1523
FT CHAIN 160 1523
FT CHAIN 161 1523
FT CHAIN 162 1523
FT CHAIN 163 1523
FT CHAIN 164 1523
FT CHAIN 165 1523
FT CHAIN 166 1523
FT CHAIN 167 1523
FT CHAIN 168 1523
FT CHAIN 169 1523
FT CHAIN 170 1523
FT CHAIN 171 1523
FT CHAIN 172 1523
FT CHAIN 173 1523
FT CHAIN 174 1523
FT CHAIN 175 1523
FT CHAIN 176 1523
FT CHAIN 177 1523
FT CHAIN 178 1523
FT CHAIN 179 1523
FT CHAIN 180 1523
FT CHAIN 181 1523
FT CHAIN 182 1523
FT CHAIN 183 1523
FT CHAIN 184 1523
FT CHAIN 185 1523
FT CHAIN 186 1523
FT CHAIN 187 1523
FT CHAIN 188 1523
FT CHAIN 189 1523
FT CHAIN 190 1523
FT CHAIN 191 1523
FT CHAIN 192 1523
FT CHAIN 193 1523
FT CHAIN 194 1523
FT CHAIN 195 1523
FT CHAIN 196 1523
FT CHAIN 197 1523
FT CHAIN 198 1523
FT CHAIN 199 1523
FT CHAIN 200 1523
FT CHAIN 201 1523
FT CHAIN 202 1523
FT CHAIN 203 1523
FT CHAIN 204 1523
FT CHAIN 205 1523
FT CHAIN 206 1523
FT CHAIN 207 1523
FT CHAIN 208 1523
FT CHAIN 209 1523
FT CHAIN 210 1523
FT CHAIN 211 1523
FT CHAIN 212 1523
FT CHAIN 213 1523
FT CHAIN 214 1523
FT CHAIN 215 1523
FT CHAIN 216 1523
FT CHAIN 217 1523
FT CHAIN 218 1523
FT CHAIN 219 1523
FT CHAIN 220 1523
FT CHAIN 221 1523
FT CHAIN 222 1523
FT CHAIN 223 1523
FT CHAIN 224 1523
FT CHAIN 225 1523
FT CHAIN 226 1523
FT CHAIN 227 1523
FT CHAIN 228 1523
FT CHAIN 229 1523
FT CHAIN 230 1523
FT CHAIN 231 1523
FT CHAIN 232 1523
FT CHAIN 233 1523
FT CHAIN 234 1523
FT CHAIN 235 1523
FT CHAIN 236 1523
FT CHAIN 237 1523
FT CHAIN 238 1523
FT CHAIN 239 1523
FT CHAIN 240 1523
FT CHAIN 241 1523
FT CHAIN 242 1523
FT CHAIN 243 1523
FT CHAIN 244 1523
FT CHAIN 245 1523
FT CHAIN 246 1523
FT CHAIN 247 1523
FT CHAIN 248 1523
FT CHAIN 249 1523
FT CHAIN 250 1523
FT CHAIN 251 1523
FT CHAIN 252 1523
FT CHAIN 253 1523
FT CHAIN 254 1523
FT CHAIN 255 1523
FT CHAIN 256 1523
FT CHAIN 257 1523
FT CHAIN 258 1523
FT CHAIN 259 1523
FT CHAIN 260 1523
FT CHAIN 261 1523
FT CHAIN 262 1523
FT CHAIN 263 1523
FT CHAIN 264 1523
FT CHAIN 265 1523
FT CHAIN 266 1523
FT CHAIN 267 1523
FT CHAIN 268 1523
FT CHAIN 269 1523
FT CHAIN 270 1523
FT CHAIN 271 1523
FT CHAIN 272 1523
FT CHAIN 273 1523
FT CHAIN 274 1523
FT CHAIN 275 1523
FT CHAIN 276 1523
FT CHAIN 277 1523
FT CHAIN 278 1523
FT CHAIN 279 1523
FT CHAIN 280 1523
FT CHAIN 281 1523
FT CHAIN 282 1523
FT CHAIN 283 1523
FT CHAIN 284 1523
FT CHAIN 285 1523
FT CHAIN 286 1523
FT CHAIN 287 1523
FT CHAIN 288 1523
FT CHAIN 289 1523
FT CHAIN 290 1523
FT CHAIN 291 1523
FT CHAIN 292 1523
FT CHAIN 293 1523
FT CHAIN 294 1523
FT CHAIN 295 1523
FT CHAIN 296 1523
FT CHAIN 297 1523
FT CHAIN 298 1523
FT CHAIN 299 1523
FT CHAIN 300 1523
FT CHAIN 301 1523
FT CHAIN 302 1523
FT CHAIN 303 1523
FT CHAIN 304 1523
FT CHAIN 305 1523
FT CHAIN 306 1523
FT CHAIN 307 1523
FT CHAIN 308 1523
FT CHAIN 309 1523
FT CHAIN 310 1523
FT CHAIN 311 1523
FT CHAIN 312 1523
FT CHAIN 313 1523
FT CHAIN 314 1523
FT CHAIN 315 1523
FT CHAIN 316 1523
FT CHAIN 317 1523
FT CHAIN 318 1523
FT CHAIN 319 1523
FT CHAIN 320 1523
FT CHAIN 321 1523
FT CHAIN 322 1523
FT CHAIN 323 1523
FT CHAIN 324 1523
FT CHAIN 325 1523
FT CHAIN 326 1523
FT CHAIN 327 1523
FT CHAIN 328 1523
FT CHAIN 329 1523
FT CHAIN 330 1523
FT CHAIN 331 1523
FT CHAIN 332 1523
FT CHAIN 333 1523
FT CHAIN 334 1523
FT CHAIN 335 1523
FT CHAIN 336 1523
FT CHAIN 337 1523
FT CHAIN 338 1523
FT CHAIN 339 1523
FT CHAIN 340 1523
FT CHAIN 341 1523
FT CHAIN 342 1523
FT CHAIN 343 1523
FT CHAIN 344 1523
FT CHAIN 345 1523
FT CHAIN 346 1523
FT CHAIN 347 1523
FT CHAIN 348 1523
FT CHAIN 349 1523
FT CHAIN 350 1523
FT CHAIN 351 1523
FT CHAIN 352 1523
FT CHAIN 353 1523
FT CHAIN 354 1523
FT CHAIN 355 1523
FT CHAIN 356 1523
FT CHAIN 357 1523
FT CHAIN 358 1523
FT CHAIN 359 1523
FT CHAIN 360 1523
FT CHAIN 361 1523
FT CHAIN 362 1523
FT CHAIN 363 1523
FT CHAIN 364 1523
FT CHAIN 365 1523
FT CHAIN 366 1523
FT CHAIN 367 1523
FT CHAIN 368 1523
FT CHAIN 369 1523
FT CHAIN 370 1523
FT CHAIN 371 1523
FT CHAIN 372 1523
FT CHAIN 373 1523
FT CHAIN 374 1523
FT CHAIN 375 1523
FT CHAIN 376 1523
FT CHAIN 377 1523
FT CHAIN 378 1523
FT CHAIN 379 1523
FT CHAIN 380 1523
FT CHAIN 381 1523
FT CHAIN 382 1523
FT CHAIN 383 1523
FT CHAIN 384 1523
FT CHAIN 385 1523
FT CHAIN 386 1523
FT CHAIN 387 1523
FT CHAIN 388 1523
FT CHAIN 389 1523
FT CHAIN 390 1523
FT CHAIN 391 1523
FT CHAIN 392 1523
FT CHAIN 393 1523
FT CHAIN 394 1523
FT CHAIN 395 1523
FT CHAIN 396 1523
FT CHAIN 397 1523
FT CHAIN 398 1523
FT CHAIN 399 1523
FT CHAIN 400 1523
FT CHAIN 401 1523
FT CHAIN 402 1523
FT CHAIN 403 1523
FT CHAIN 404 1523
FT CHAIN 405 1523
FT CHAIN 406 1523
FT CHAIN 407 1523
FT CHAIN 408 1523
FT CHAIN 409 1523
FT CHAIN 410 1523
FT CHAIN 411 1523
FT CHAIN 412 1523
FT CHAIN 413 1523
FT CHAIN 414 1523
FT CHAIN 415 1523
FT CHAIN 416 1523
FT CHAIN 417 1523
FT CHAIN 418 1523
FT CHAIN 419 1523
FT CHAIN 420 1523
FT CHAIN 421 1523
FT CHAIN 422 1523
FT CHAIN 423 1523
FT CHAIN 424 1523
FT CHAIN 425 1523
FT CHAIN 426 1523
FT CHAIN 427 1523
FT CHAIN 428 1523
FT CHAIN 429 1523
FT CHAIN 430 1523
FT CHAIN 431 1523
FT CHAIN 432 1523
FT CHAIN 433 1523
FT CHAIN 434 1523
FT CHAIN 435 1523
FT CHAIN 436 1523
FT CHAIN 437 1523
FT CHAIN 438 1523
FT CHAIN 439 1523
FT CHAIN 440 1523
FT CHAIN 441 1523
FT CHAIN 442 1523
FT CHAIN 443 1523
FT CHAIN 444 1523
FT CHAIN 445 1523
FT CHAIN 446 1523
FT CHAIN 447 1523
FT CHAIN 448 1523
FT CHAIN 449 1523
FT CHAIN 450 1523
FT CHAIN 451 1523
FT CHAIN 452 1523
FT CHAIN 453 1523
FT CHAIN 454 1523
FT CHAIN 455 1523
FT CHAIN 456 1523
FT CHAIN 457 1523
FT CHAIN 458 1523
FT CHAIN 459 1523
FT CHAIN 460 1523
FT CHAIN 461 1523
FT CHAIN 462 1523
FT CHAIN 463 1523
FT CHAIN 464 1523
FT CHAIN 465 1523
FT CHAIN 466 1523
FT CHAIN 467 1523
FT CHAIN 468 1523
FT CHAIN 469 1523
FT CHAIN 470 1523
FT CHAIN 471 1523
FT CHAIN 472 1523
FT CHAIN 473 1523
FT CHAIN 474 1523
FT CHAIN 475 1523
FT CHAIN 476 1523
FT CHAIN 477 1523
FT CHAIN 478 1523
FT CHAIN 479 1523
FT CHAIN 480 1523
FT CHAIN 481 1523
FT CHAIN 482 1523
FT CHAIN 483 1523
FT CHAIN 484 1523
FT CHAIN 485 1523
FT CHAIN 486 1523
FT CHAIN 487 1523
FT CHAIN 488 1523
FT CHAIN 489 1523
FT CHAIN 490 1523
FT CHAIN 491 1523
FT CHAIN 492 1523
FT CHAIN 493 1523
FT CHAIN 494 1523
FT CHAIN 495 1523
FT CHAIN 496 1523
FT CHAIN 497 1523
FT CHAIN 498 1523
FT CHAIN 499 1523
FT CHAIN 500 1523
FT CHAIN 501 1523
FT CHAIN 502 1523
FT CHAIN 503 1523
FT CHAIN 504 1523
FT CHAIN 505 1523
FT CHAIN 506 1523
FT CHAIN 507 1523
FT CHAIN 508 1523
FT CHAIN 509 1523
FT CHAIN 510 1523
FT CHAIN 511 1523
FT CHAIN 512 1523
FT CHAIN 513 1523
FT CHAIN 514 1523
FT CHAIN 515 1523
FT CHAIN 516 1523
FT CHAIN 517 1523
FT CHAIN 518 1523
FT CHAIN 519 1523
FT CHAIN 520 1523
FT CHAIN 521 1523
FT CHAIN 522 1523
FT CHAIN 523 1523
FT CHAIN 524 1523
FT CHAIN 525 1523
FT CHAIN 526 1523
FT CHAIN 527 1523
FT CHAIN 528 1523
FT CHAIN 529 1523
FT CHAIN 530 1523
FT CHAIN 531 1523
FT CHAIN 532 1523
FT CHAIN 533 1523
FT CHAIN 534 1523
FT CHAIN 535 1523
FT CHAIN 536 1523
FT CHAIN 537 1523
FT CHAIN 538 1523
FT CHAIN 539 1523
FT CHAIN 540 1523
FT CHAIN 541 1523
FT CHAIN 542 1523
FT CHAIN 543 1523
FT CHAIN 544 1523
FT CHAIN 545 1523
FT CHAIN 546 1523
FT CHAIN 547 1523
FT CHAIN 548 1523
FT CHAIN 549 1523
FT CHAIN 550 1523
FT CHAIN 551 1523
FT CHAIN 552 1523
FT CHAIN 553 1523
FT CHAIN 554 1523
FT CHAIN 555 1523
FT CHAIN 556 1523
FT CHAIN 557 1523
FT CHAIN 558 1523
FT CHAIN 559 1523
FT CHAIN 560 1523
FT CHAIN 561 1523
FT CHAIN 562 1523
FT CHAIN 563 1523
FT CHAIN 564 1523
FT CHAIN 565 1523
FT CHAIN 566 1523
FT CHAIN 567 1523
FT CHAIN 568 1523
FT CHAIN 569 1523
FT CHAIN 570 1523
FT CHAIN 571 1523
FT CHAIN 572 1523
FT CHAIN 573 1523
FT CHAIN 574 1523
FT CHAIN 575 1523
FT CHAIN 576 1523
FT CHAIN 577 1523
FT CHAIN 578 1523
FT CHAIN 579 1523
FT CHAIN 580 1523
FT CHAIN 581 1523
FT CHAIN 582 1523
FT CHAIN 583 1523
FT CHAIN 584 1523
FT CHAIN 585 1523
FT CHAIN 586 1523
FT CHAIN 587 1523
FT CHAIN 588 1523
FT CHAIN 589 1523
FT CHAIN 590 1523
FT CHAIN 591 1523
FT CHAIN 592 1523
FT CHAIN 593 1523
FT CHAIN 594 1523
FT CHAIN 595 1523
FT CHAIN 596 1523
FT CHAIN 597 1523
FT CHAIN 598 1523
FT CHAIN 599 1523
FT CHAIN 600 1523
FT CHAIN 601 1523
FT CHAIN 602 1523
FT CHAIN 603 1523
FT CHAIN 604 1523
FT CHAIN 605 1523
FT CHAIN 606 1523
FT CHAIN 607 1523
FT CHAIN 608 1523
FT CHAIN 609 1523
FT CHAIN 610 1523
FT CHAIN 611 1523
FT CHAIN 612 1523
FT CHAIN 613 1523
FT CHAIN 614 1523
FT CHAIN 615 1523
FT CHAIN 616 1523
FT CHAIN 617 1523
FT CHAIN 618 1523
FT CHAIN 619 1523
FT CHAIN 620 1523
FT CHAIN 621 1523
FT CHAIN 622 1523
FT CHAIN 623 1523
FT CHAIN 624 1523
FT CHAIN 625 1523
FT CHAIN 626 1523
FT CHAIN 627 1523
FT CHAIN 628 1523
FT CHAIN 629 1523
FT CHAIN 630 1523
FT CHAIN 631 1523
FT CHAIN 632 1523
FT CHAIN 633 1523
FT CHAIN 634 1523
FT CHAIN 635 1523
FT CHAIN 636 1523
FT CHAIN 637 1523
FT CHAIN 638 1523
FT CHAIN 639 1523
FT CHAIN 640 1523
FT CHAIN 641 1523
FT CHAIN 642 1523
FT CHAIN 643 1523
FT CHAIN 644 1523
FT CHAIN 645 1523
FT CHAIN 646 1523
FT CHAIN 647 1523
FT CHAIN 648 1523
FT CHAIN 649 1523
FT CHAIN 650 1523
FT CHAIN 651 1523
FT CHAIN 652 1523
FT CHAIN 653 1523
FT CHAIN 654 1523
FT CHAIN 655 1523
FT CHAIN 656 1523
FT CHAIN 657 1523
FT CHAIN 658 1523
FT CHAIN 659 1523
FT CHAIN 660 1523
FT CHAIN 661 1523
FT CHAIN 662 1523
FT CHAIN 663 1523
FT CHAIN 664 1523
FT CHAIN 665 1523
FT CHAIN 666 1523
FT CHAIN 667 1523
FT CHAIN 668 1523
FT CHAIN 669 1523
FT CHAIN 670 1523
FT CHAIN 671 1523
FT CHAIN 672 1523
FT CHAIN 673 1523
FT CHAIN 674 1523
FT CHAIN 675 1523
FT CHAIN 676 1523
FT CHAIN 677 1523
FT CHAIN 678 1523
FT CHAIN 679 1523
FT CHAIN 680 1523
FT CHAIN 681 1523
FT CHAIN 682 1523
FT CHAIN 683 1523
FT CHAIN 684 1523
FT CHAIN 685 1523
FT CHAIN 686 1523
FT CHAIN 687 1523
FT CHAIN 688 1523
FT CHAIN 689 1523
FT CHAIN 690 1523
FT CHAIN 691 1523
FT CHAIN 692 1523
FT CHAIN 693 1523
FT CHAIN 694 1523
FT CHAIN 695 1523
FT CHAIN 696 1523
FT CHAIN 697 1523
FT CHAIN 698 1523
FT CHAIN 699 1523
FT CHAIN 700 1523
FT CHAIN 701 1523
FT CHAIN 702 1523
FT CHAIN 703 1523
FT CHAIN 704 1523
FT CHAIN 705 1523
FT CHAIN 706 1523
FT CHAIN 707 1523
FT CHAIN 708 1523
FT CHAIN 709 1523
FT CHAIN 710 1523
FT CHAIN 711 1523
FT CHAIN 712 1523
FT CHAIN 713 1523
FT CHAIN 714 1523
FT CHAIN 715 1523
FT CHAIN 716 1523
FT CHAIN 717 1523
FT CHAIN 718 1523
FT CHAIN 719 1523
FT CHAIN 720 1523
FT CHAIN 721 1523
FT CHAIN 722 1523
FT CHAIN 723 1523
FT CHAIN 724 1523
FT CHAIN 725 1523
FT CHAIN 726 1523
FT CHAIN 727 1523
FT CHAIN 728 1523
FT CHAIN 729 1523
FT CHAIN 730 1523
FT CHAIN 731 1523
FT CHAIN 732 1523
FT CHAIN 733 1523
FT CHAIN 734 1523
FT CHAIN 735 1523
FT CHAIN 736 1523
FT CHAIN 737 1523
FT CHAIN 738 1523
FT CHAIN 739 1523
FT CHAIN 740 1523
FT CHAIN 741 1523
FT CHAIN 742 1523
FT CHAIN 743 1523
FT CHAIN 744 1523
FT CHAIN 745 1523
FT CHAIN 746 1523
FT CHAIN 747 1523
FT CHAIN 748 1523
FT CHAIN 749 1523
FT CHAIN 750 1523
FT CHAIN 751 1523
FT CHAIN 752 1523
FT CHAIN 753 1523
FT CHAIN 754 1523
FT CHAIN 755 1523
FT CHAIN 756 1523
FT CHAIN 757 1523
FT CHAIN 758 1523
FT CHAIN 759 1523
FT CHAIN 760 1523
FT CHAIN 761 1523
FT CHAIN 762 1523
FT CHAIN 763 1523
FT CHAIN 764 1523
FT CHAIN 765 1523
FT CHAIN 766 1523
FT CHAIN 767 1523
FT CHAIN 768 1523
FT CHAIN 769 1523
FT CHAIN 770 1523
FT CHAIN 771 1523
FT CHAIN 772 1523
FT CHAIN 773 1523
FT CHAIN 774 1523
FT CHAIN 775 1523
FT CHAIN 776 1523
FT

```

Db 738 KTEYKGVYDLTEGPFYFANGILTNNSLPSIIISHVSPPTLNREG--CEEYD--E 793
QY 472 APEYKHFRRCPFPFPKTVLERLELRVAEMKKFPPDSEPRLLDEROKALKVLANA 531
Db 794 APQVGHFFCKDPFQFISLLDLDEROKVAKHKA--TVDEIEKKLLDYRORAIKIANS 852
QY 532 SYGFMGSGARWYCRECKAKATANGRHLIRTAI-NIARKLAKIY-----GDDP----- 580
Db 853 FYGIYGAAKARWYCKECSATANGROYIETTKIEIEKFGKVLADSVTGDEIVTIRR 912
QY 581 ----- 580
Db 913 NGRIEFPVIEKLFERVDHVRGEKEYCVLGVEALTLDRGRLLWKKVPPYWRHKTDKRIY 972
QY 581 ----- 580
Db 973 RWFTNSWYLDVTEDEHSLIGYLTNSKVKPGKPLKRLVEVKEPELGGKYSLLTPNPIA 1032
QY 581 ----- 580
1033 RTIQAAPLAVLMLIGLIGVDGKMGSGSNKAYVGLSCGLDAEIERKVLNPLREASV 1092
QY 581 ----- 580
1093 ISNYDKSKGVDVSLSKMLAGFMVKYFKDENGKKAIPSEFENLPREYIEAFAGLPSAD 1152
QY 581 ----- 583
1153 GTVALRGIPERILTVSNRELSDAVRKLMLVGVNSLFTKPNRYLEKESGTHSIHR 1212
QY 584 ----- 587
1213 IKNHGRPADRIGFLIDRKSTKLSENLGCHNKKRAKYKDFPLVYPRKIEETTYGGYVDI 1272
QY 588 ----- 612
1273 EVEGTHRFANGILVHNTDGFATIPGADAFYVKKKAEPFLNTPKPLGELLEBEYGF 1332
QY 613 KRLPFTAKKRYAGLDEGRIDIVGEFVNRDGCBLAEVOTKVEIVLKTSEVNAVEY 672
Db 1333 RRGFFV--TKKKYAVIDEKGKLTTRGLEIVRRDWSVAKEQARVLEALIRHGDVEAARI 1391
QY 673 VRKIVKELEBEGVPIEKVIMKTLISKRLSEYTTTEAPHVAAKMLNSGVRSPDKIGYV 732
Db 1392 VKETEKLSKVEPEKLVHIEQITRELDKDYKATGPHVAAIKRLAARGIKVRPGTIVSYI 1451
733 IVKGGRIISORAWPYFWVKDPS--QIDVTYVVDHQQIIPALRLIIGYFITEKTKLKASATG 790
1452 VLKSGRIGDRTIP--PDEFDPTKRHYDAEYIENQVLPAYERILKAPYKKEDELRYQKTR 1510
QY 791 QKITFDPLAKSK 803
Db 1511 QVGLGAWLKMGKK 1523

RESULT 14
DPOOL_PYRO STANDARD; PRT; 1671 AA.
AC P77933;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE DNA polymerase (EC 2.7.7.7) [Contains: Endonuclease PI-PKOI
DE (EC 3.1.1.1) (Pko pol-1 intein) (IVS-A); Endonuclease PI-PKOI
DE (EC 3.1.1.1) (Pko pol-2 intein) (IVS-B)].
OS Pyrococcus kodakarensis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_taxid=69014;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

```

```

RC STRAIN=KOD1.
RX MEDLINE=98027387; PubMed=9361436;
RA Takagi M., Nishioka M., Kakihara H., Kitabayashi M., Inoue H.,
RA Kawakami B., Oka M., Imanaka T.;
RT "Characterization of DNA polymerase from Pyrococcus sp. strain KOD1
RT and its application to PCR."
RL Appl. Environ. Microbiol. 63:4504-4510(1997).
RN [2].
RP CHARACTERIZATION OF INTEIN.
RC STRAIN=KOD1;
RX MEDLINE=98416198; PubMed=9742242;
RA Nishioka M., Fujiwara S., Takagi M., Imanaka T.;
RT "Characterization of two intein homing endonucleases encoded in the
RT DNA polymerase gene of Pyrococcus kodakarensis strain KOD1."
RL Nucleic Acids Res. 26:4409-4412(1998).
CC -1- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
CC INTEIN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY
CC ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE
CC INTEIN. PI-PKOI RECOGNIZES 5'-GATTAGATCCCTACC-3' AND PI-PROII
CC RECOGNIZES 5'-CAGCTACTACGGTTAC-3'. BOTH ARE THERMOSTABLE.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -1- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING
CC ENDONUCLEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D29671; BAA06142.2; -.
DR HSP; P56689; ITGO.
DR REBASE; 3792; PI-PKOI.
DR REBASE; 3793; PI-PKOII.
DR InterPro; IPR002064; DNA_pol_B.
DR InterPro; IPR003586; Hedgehog_hinc.
DR InterPro; IPR003587; Hedgehog_hinc.
DR InterPro; IPR002203; InteIn.
DR InterPro; IPR004042; InteIn_endonuc.
DR InterPro; IPR004578; Pol2.
DR Pfam; PF00136; DNA_pol_B_3.
DR Pfam; PF03104; DNA_pol_B_exo_1.
DR PRINTS; PR00379; INTEIN.
DR SMART; SM00305; HincC_2.
DR SMART; SM00306; HincC_2.
DR SMART; SM00486; POLBc_1.
DR TIGRFAMs; TIGR00592; pol2.2.
DR PROSITE; PS00116; DNA_POLYMERASE_B_1.
DR PROSITE; PS00818; INTEIN_C_TER_2.
DR PROSITE; PS00819; INTEIN_ENDONUCLEASE_2.
DR PROSITE; PS00817; INTEIN_N_TER_2.
KM Transferase; DNA-directed DNA polymerase; DNA replication;
KM DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
KW Protein splicing; Intron homing.
FT CHAIN 1 406 DNA POLYMERASE, 1ST PART.
FT CHAIN 407 766 ENDONUCLEASE PI-PKO I.
FT CHAIN 767 851 DNA POLYMERASE, 2ND PART.
FT CHAIN 852 1388 ENDONUCLEASE PI-PKO II.
FT CHAIN 1389 1671 DNA POLYMERASE, 3RD PART.
SQ SEQUENCE 1671 AA; 193490 MW; 4A17F1C8120E8455 CRC64;

Query Match 18.5%; Score 776; DB 1; Length 1671;
Best Local Similarity 17.6%; Pred. No. 1e-41;
Matches 304; Conservative 145; Mismatches 287; Indels 990; Gaps 24;
QY 8 VLDSYVAVKEQVVIWGIAENGERVVLIDRSFRFYFALLAPGADPKV-----AQRIR 63

```

```

Db      2  ILDTDYTEDGKPVIRIFK-KENGEFKIEYDRFFEEFYALLKODSAIEEVKKITMERBG 60
Qy      64  ALSRPSPLIGVEDDKKCYFGRPRVLRITVLPEVAREYELKANDGVLEDITF 123
Db      61  TVTVYGR-----VEKQKQKFLGRVYEVKVLFTTHPOVPAIRDKIREHPANIDYEDIDIF 116
Qy      124  AMXYLIDHDLFPFTWYRVEAPLENNKMGFVNDKTVLVSREPLYGALAPTLPLDRL 183
Db      117  AKKYLIDKGLV-----PMEGD-----BELKML 138
Qy      184  APDIEVYSKQSPREPERDPIVIAKTDGDEVL-----FLAEGKDRKRPFEVY 235
Db      139  APDIEVYHEGEEFAE-GPILMISYADEGCAVITWKVNDLPYVDVSTEREMIKKFLV 197
Qy      236  VKRYDPIIVGYNHNPMPYLLRRARILGIKLDTVRVGAEPPTSVGH---VSYPGL 292
Db      198  VKKEDVDVLTYNQDNDFAYLKKECKGINFALGRD-GSEPKIQHMGDRFAVEYKGR 256
Qy      293  NVLDYDAEEMPEIKIKSLAEVAYLGVMKKSEFVIMMWEIPYWDOPKKRPLLOVAR 352
Db      257  HFDLYVIRRTINLPYTLLEAVYEAVGQPK-EKVYAE--EITTAMETGENLERVAR 313
Qy      353  DQVATYGLAEKILPFAIQSYVTGLPDQVAMSVGRLEKYLIRAPFKKELVPRV 412
Db      314  EDKATYELGKREFLPMEQSLRLLGOSLMVSRSSITGVLVEPLIRKQYERNEBLANK 373
Qy      413  RPE-----ETYGALVLEPLAGVHENAVLDF----- 439
Db      374  EKELARRQSTIEGKYVKEPERGEMENLVYDFRCHPADTKVYVKGKGINISEVQGDYV 433
Qy      440  ----- 439
Db      434  LGIDGQVRKVMYDYKGLVNINGLKCTPNHKLPVYTKOERQTRIRDSIAKSLTKKY 493
Qy      440  ----- 439
Db      494  KGIITPLPYEIGRATSENIPEBEVLKGLAGIILAEGLLRKQVDFDSRKKRISH 553
Qy      440  ----- 439
Db      554  QYRVEITIGKDEEFRRIRITYFERLFGITPISSEKGTNAVTLKAKQVYLKVEIMD 613
Qy      440  ----- 439
Db      614  NIESLHAPSVLRGFFEGDSVNRVRSIVATGCTKNEKIKLVSKLSQGLGIDHYTYQ 673
Qy      440  ----- 439
Db      674  YQENGKDRRYILEITGKGLILFQTLIGISERKALLNKAISQREMANLENNGFYRLS 733
Qy      440  ----- 467
Db      734  EFNVSTYEYEGKVNDLLEGTPTYFANGILTHNSLYPSTITTHNVSPDTLNR----- 786
Qy      468  GCNE--APEVGRFRRCPPGFVTLERLLERLKRKRAEMKKYPPSPFYRLIDRQKA 524
Db      787  GCKEYDVAPOVGRPCPKPGFIPSLIGLLEEROKIKKKKA--TIDPIERKILDYROA 845
Qy      525  LKVL----- 529
Db      846  IKILANSILPEEWLPLVEEGEVHFRVIGELIDRMEENAGVKGEGETEVLVSGLEVPS 905
Qy      530  ----- 529
Db      906  FNRRTNKAELKRVKALLRHDSQKVYTIIRLKSGRKIKITSGHSLFSVRNGELVEYVGD 965
Qy      530  ----- 529
Db      966  KPEDVAVVRRLLEPERNHVLNVLLELGTPEEETLDIVMTIPVKGKKNFKQMLATLRM 1025
Qy      530  ----- 529

```

```

Db      1026  IFGEERPRPTARVLRHLEDLGYYRLKKIGYELDMDSLKMYRRLYEALVENRYNNGKR 1085
Qy      530  ----- 529
Db      1086  EYLVENSTRDVAVGIMPLKELKEMKIGTLNGFRMKLEBVDESLAKLGLYVSEGYARQ 1145
Qy      530  ----- 529
Db      1146  RNPANGMSYVKLYNEDEVDLDMERLASRFGKVRGRNGYVEIPKKIGYLLFENNCGVL 1205
Qy      530  ----- 529
Db      1206  AENKRIPEFVTPKGVLAFLLEGYFIGDQDVHPNKRRLSTKSELLANQVLILNSGV 1265
Qy      530  ----- 529
Db      1266  SAVLGHDSGVRYVINEELPFVKLDKKNAYYSHVPEVILSEVFGKVFQKVPQTR 1325
Qy      530  ----- 529
Db      1326  KWEEDRLDPEKQRLSLIEGDVLDREVSVDEYDGYVYDLSVEDNENFLVGRGLY 1385
Qy      530  ----- 586
Db      1386  AHNSTYGYIVARAHMYCKEASVTAMGREYITMTIKEIERKGFVITSDDGFFAT 1445
Qy      587  ---DPEKVE---NFIKTIKEEL--GEIKLEKVRRLPFTBAKKRYAGLLEDGIDIVG 637
Db      1446  PGADAETVKKAMEFLKYINKLPGALELEYBGFYKRGFPV-TKKRYAVIDEGKLTITG 1504
Qy      638  FEAVRGDMCELAKEVQTVKVEIVLKTSEVNAKAVEYRKIVLEBEGVPLEKLVIKTLS 697
Db      1505  LEIVRDMSEIAKEVQAVLEALLKQGDVEKVRIVKVEYTEKLSKRYEPEKLVIEHQIT 1564
Qy      698  KRLEEYTPAPHVNAKMLLSAGYRVSPODKIGYIVVKGGRISQRPMPFMKDPDS--Q 755
Db      1565  RDKDYKATGPVAVAKKLAARGVKIRPGTVISYIVLKGSGRIDRAIP-FDFEDPTKX 1623
Qy      756  IDVTVYVHDQIIPALRILIGYFGITEKCLKASATGOKTLFPFLAK 801
Db      1624  YDAEYIENQVPAVERILRAFGYRKEDLRKYQTKQVGLSMLMPK 1669

```

```

RESUT. 15
DPD SCHPO STANDARD; PRT; 1086 AA.
AC P30316; Q10016; Q9USU0;
DT 01-APR-1993 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase delta catalytic subunit (EC 2.7.7.7) (DNA polymerase
DE III).
OS POL3 OR POLD OR SPBC36.04.
GN Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93184400; PubMed=8443413;
RA Plagnade G., Bouvier D., de Recondo A.M., Baldacci G.;
RT "Characterization of the Pol3 gene product from Schizosaccharomycetes
RT pombe indicates inter-species conservation of the catalytic subunit
RT of DNA polymerase delta."
RL J. Mol. Biol. 222:209-218(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93184400; PubMed=8443413;
RA Pak H., Francesconi S., Wang T.S.F.;
RT "Cell cycle expression of two replicative DNA polymerases alpha and
RT delta from Schizosaccharomycetes pombe."
RL Mol. Biol. Cell 4:145-157(1993).

```


Wed Apr 23 08:09:18 2003

Job time : 28 secs

us-10-034-849-2.rsp

Page 16

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:28:09 ; Search time 41 Seconds
(without alignments)
4035.508 Million cell updates/sec

Title: US-10-034-849-2
Perfect score: 4188
Sequence: 1 MTEVFVTVLDSSEYEVGKEP.....LKASATGQKTLFPFLAKSK 803

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL_21:*

- 1: sp archaea:*
- 2: sp bacteria:*
- 3: sp fungi:*
- 4: sp human:*
- 5: sp invertebrate:*
- 6: sp mammal:*
- 7: sp mhc:*
- 8: sp organelle:*
- 9: sp phage:*
- 10: sp plant:*
- 11: sp rodent:*
- 12: sp virus:*
- 13: sp vertebrate:*
- 14: sp unclassified:*
- 15: sp virus:*
- 16: sp bacteriophage:*
- 17: sp archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3033.5	72.4	803	1 Q59691	Q59691 pyrodicticum
2	2248	53.7	781	1 Q96VY1	Q96VY1 sulfobolus
3	2246	53.6	781	1 Q9P9N1	Q9P9N1 sulfobolus
4	2156	51.5	785	1 Q8ZVW2	Q8ZVW2 pyrobaculum
5	2129.5	50.8	785	1 Q9P9M2	Q9P9M2 pyrobaculum
6	1268	30.2	775	1 Q9P9K4	Q9P9K4 pyrobaculum
7	1264	30.2	775	1 Q9H06	Q9H06 pyrobaculum
8	1220	29.1	830	1 Q8TWJ5	Q8TWJ5 methanopyru
9	1208	28.8	759	1 Q9H93	Q9H93 pyrococcus
10	852.5	20.4	937	1 Q8TSB3	Q8TSB3 methanopyru
11	760.5	18.2	1107	4 Q96H8	Q96H8 homo sapien
12	743.5	17.8	1105	11 Q91VT0	Q91VT0 mus sapien
13	737.5	17.6	1104	3 Q8X07	Q8X07 neurospora
14	732.5	17.5	974	5 Q8SQS5	Q8SQS5 encephalito
15	718.5	17.2	901	17 Q9HRV9	Q9HRV9 halobacteri
16	655.5	15.7	855	17 Q8ZVP9	Q8ZVP9 pyrobaculum

17	649.5	15.5	914	1 Q59690	Q59690 pyrodicticum
18	648	15.5	876	17 Q971C7	Q971C7 sulfobolus
19	646.5	15.4	796	17 Q9H06	Q9H06 thermoplasm
20	622	14.9	800	17 Q97AH3	Q97AH3 thermoplasm
21	575.5	13.7	1458	13 Q9DE46	Q9DE46 xenopus lae
22	569	13.6	787	16 Q9KSP4	Q9KSP4 vibrio chol
23	564.5	13.5	844	1 Q31096	Q31096 cenarchaeum
24	560.5	13.4	1016	12 Q993K6	Q993K6 callitrichi
25	552	13.2	845	1 Q74046	Q74046 cenarchaeum
26	540	12.9	1492	10 Q9FHA3	Q9FHA3 arabidopsis
27	539	12.9	2909	4 Q9NU25	Q9NU25 homo sapien
28	536	12.8	1015	12 Q8UZD7	Q8UZD7 cercopithec
29	527.5	12.6	787	16 Q91ZL1	Q91ZL1 pseudomonas
30	524	12.5	1013	12 Q91IX9	Q91IX9 retroperito
31	521.5	12.5	1013	12 Q41208	Q41208 pseudorabito
32	515	12.3	1048	12 Q63390	Q63390 pseudorabito
33	511	12.2	956	5 Q8SRW2	Q8SRW2 encephalito
34	511	12.2	1026	12 Q36363	Q36363 alcelaphine
35	510	12.2	1004	12 Q90628	Q90628 porcine lym
36	507.5	12.1	1443	5 Q9NKH1	Q9NKH1 caenorhabdi
37	506.5	12.1	1012	12 Q40915	Q40915 kaposi's sa
38	506.5	12.1	1012	12 Q40910	Q40910 kaposi's sa
39	506.5	12.1	1472	5 Q9VD90	Q9VD90 drosophila
40	506	12.1	774	2 Q9FI75	Q9FI75 pseudomonas
41	503	12.0	1041	12 Q9PKT8	Q9PKT8 elephant he
42	501	12.0	1005	12 Q99D20	Q99D20 bovine hept
43	501	12.0	1009	12 Q9YUQ4	Q9YUQ4 atelina her
44	499.5	11.9	789	16 Q8ZIH9	Q8ZIH9 yersinia pe
45	497	11.9	1415	5 Q77034	Q77034 drosophila

ALIGNMENTS

RESULT 1

ID Q59691 PRELIMINARY; PRT; 803 AA.

AC Q59691; 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE DNA polymerase (EC 2.7.7.7).

GN POLB

OS Pyrodicticum occultum.

OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;

OC Pyrodictiaceae; Pyrodicticum.

OX NCBI_Taxid=2309;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DSM2709;

RX MEDLINE=95238290; PubMed=721707;

RA Demori T., Ishino Y., Doi H., Kato I.;

RT "The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-like DNA polymerases";

RT J. Bacteriol. 177:2164-2177(1995).

FL J. Bacteriol. 177:2164-2177(1995).

CC - CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE + {DNA} (N).

CC - SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

CC EMBL; D38574; BAA07580.1; -.

DR HSSP; P56689; ITGO.

DR InterPro; IPR002064; DNA_pol_B.

DR InterPro; IPR004578; Pol2.

DR Pfam; PF00136; DNA_pol_B_1.

DR Pfam; PF03104; DNA_pol_B_exo; 1.

DR PRINTS; PR00106; DNAPOLB.

DR SMART; SM00486; POLBc; 1.

DR TIGRFAMs; TIGR00592; pol2; 1.

DR PROSITE; PS00116; DNA_POLYMERASE_B_1.

DR DNA replication; DNA-binding; DNA-directed DNA polymerase;

DR Nucleotidyltransferase; Transferase.

SW SQUENCH 803 AA; 92656 MW; E9C092F26ABD23FA CRC64;

Query Match 72.4%; Score 3033.5; DB 1; Length 803;

Best Local Similarity 70.9%; Pred. No. 6,86-193;
Matches 571; Conservative 106; Mismatches 121; Indels 7; Gaps 5;

```

QY 1 MTEVY-FTVLDSSEYVNGKEPOVIIMGIAENGERVVLIDRSFRPFYALLAPGADP--KO 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MTEIEFVLDSSEYVNGKEPOVIIMGIAENGERVVLIDRSFRPFYALLAPGADP--KO 60
QY 58 VAGRTAASRPKSPPIGVEDDKKRCRRPRVLRITVLPVAVREYELVAVDQVEDLV 117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 IASIRLSVYKSPILIDAKPDKKIFRRPKAKVITITMIPSVIRHTEBAVKIEGVDEL 120
QY 118 EADIRFANRYLIDHDLFFFTYVREAEPLNNKGRPVDDKVLYVSRPEPLYGALAPTKL 177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 EADIRFANRYLIDHDLFFFTYVREAEPLNNKGRPVDDKVLYVSRPEPLYGALAPTKL 178
QY 178 PDLRLIADPIEVYKSGSPRRPDPVIVIAVKTDDDEVLFIAEGKDRKPIREFEYVK 237
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 PWRVLAADPIEVYKSGSPRRPDPVIVIAVKTDDDEVLFIAEGKDRKPIREFEYVK 238
QY 238 RYDPDIIVGYNNNHDFMPEYLLRRARIGILDTVRVGAEPFTSVGHVGSVPRLAVDLY 297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 AFDPOIIVGYNNNHDFMPEYLLRRARIGILDTVRVGAEPFTSVGHVGSVPRLAVDLY 298
QY 298 DVAEEMPEIKIKSLEEVYELGVWKKSERVVIINMEIDPYMDDEKRPILLOYARDVNA 357
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 DVAEEMPEIKIKSLEEVYELGVWKKSERVVIINMEIDPYMDDEKRPILLOYARDVNA 358
QY 358 TVGLAEKILPRLIOLSVYTGILPDQVANSVGRLEWYLIRARFKMKELVPRNVERPEET 417
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 TVGLAEKILPRLIOLSVYTGILPDQVANSVGRLEWYLIRARFKMKELVPRNVERPEET 418
QY 418 YRGALVLEPLRGVHENIAVDPSGMPINIKINVGPTLVPRGKSGEC--GCWEAPEVK 476
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 YRGALVLEPLRGVHENIAVDPSGMPINIKINVGPTLVPRGKSGEC--GCWEAPEVK 478
QY 477 HRRRCPPGFFKTVLERLLELRKVRRAEMKKYPPDSPEYRLIDEROKALKVLANASYGM 536
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 479 HRRRCPPGFFKTVLERLLELRKVRRAEMKKYPPDSPEYRLIDEROKALKVLANASYGM 538
QY 537 GMSGARWYCRECAKAVTAMGHHILRPTAINIARLGLKVIYGDTSLSFTYDPRKVNPFK 596
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 539 GMSGARWYCRECAKAVTAMGHHILRPTAINIARLGLKVIYGDTSLSFTYDPRKVNPFK 598
QY 597 IIEELGFEIKLEKVKYKRLPFTBAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVOTV 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 599 IIEELGFEIKLEKVKYKRLPFTBAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVOTV 658
QY 657 VEIVLKTSEVNAVEYRKVYKLEBEGKPIELVIMKTLISKRLSEETTEAPRVAAAKM 716
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 659 VEIVLKTSEVNAVEYRKVYKLEBEGKPIELVIMKTLISKRLSEETTEAPRVAAAKM 718
QY 717 LSAGRVSPGDKIGVYKGGGRISQAMPYEVNDQSDIVTYVVDHIIIPALRIILGY 776
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 719 LSAGRVSPGDKIGVYKGGGRISQAMPYEVNDQSDIVTYVVDHIIIPALRIILGY 777
QY 777 FGITTEKCLKASATGOKTLFDFLAKK 801
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 778 FGITTEKCLKASATGOKTLFDFLAKK 802

```

RESULT 2

Q96YV1

ID Q96YV1

PRELIMINARY;

PRT; 781 AA.

AC Q96YV1

DT 01-DEC-2001 (Tremblere) 19, Created)

DT 01-DEC-2001 (Tremblere) 19, Last sequence update)

DT 01-JUN-2002 (Tremblere) 21, Last annotation update)

DE Putative DNA-directed DNA polymerase II.

GN ST2076.

OS Sulfolobus tokodaii.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_taxid=111955;

```

RN [1]
RP SEQUENCE FROM N. A.
RC SPTAIN=JCM 10545 / 71
RX PubMed=11572479;
RA Kawabuchi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hoshizawa T., Kato Y.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Koshida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000988; BAB67175.1; -.
DR InterPro; IPR002064; DNA_pol_B.
DR Pfam; PF00136; DNA_pol_B; 1.
DR TIGRFAMs; TIGR00592; pol2; 1.
DR PROSITE; PS00116; DNA polymerase; Hypothetical protein; Complete proteome.
KW DNA-directed DNA polymerase; Hypothetical protein; Complete proteome.
SQ
SEQUENCE 781 AA, 90447 MW, 2561383DP903ER CRC64;

```

Query Match 53.7%; Score 2248; DB 17; Length 781;
Best Local Similarity 54.0%; Pred. No. 9,4e-141;
Matches 430; Conservative 143; Mismatches 203; Indels 20; Gaps 10;

```

QY 6 FTVLDSSEYVNGKEPOVIIMGIAENGERVVLIDRSFRPFYALLAPGADPQVAAQRIRAL 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5 FTVLDSSEYVNGKEPOVIIMGIAENGERVVLIDRSFRPFYALLAPGADPQVAAQRIRAL 64
QY 66 SRKSPPIGVEDDKKRYGRRVLRITVLPVAVREYELVAVDQVEDLVLENDIRFAM 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 SRKSPPIGVEDDKKRYGRRVLRITVLPVAVREYELVAVDQVEDLVLENDIRFAM 124
QY 126 RYLDHDLFFFTYVREAEPLNNKGRPVDDKVLYVSRPEPLYGALAPTKLPLRLIAP 185
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 RYLDHDLFFFTYVREAEPLNNKGRPVDDKVLYVSRPEPLYGALAPTKLPLRLIAP 177
QY 186 DIEVYKSGSPRRPDPVIVIAVKTDDDEVLFIAEGKDRKPIREFEYVKYVDDPIIV 245
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 187 DIEVYKSGSPRRPDPVIVIAVKTDDDEVLFIAEGKDRKPIREFEYVKYVDDPIIV 244
QY 246 GYNNHPPMPYLLRRAARLIGILDTVRVGAEPFTSVGHVGSVPRLAVDLYAEMPE 305
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 247 GYNNHPPMPYLLRRAARLIGILDTVRVGAEPFTSVGHVGSVPRLAVDLYAEMPE 306
QY 306 IKIKSLEEVYELGVWKKSERVVIINMEIDPYMDDEKRPILLOYARDVNAIYGLAEKI 365
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 307 IKIKSLEEVYELGVWKKSERVVIINMEIDPYMDDEKRPILLOYARDVNAIYGLAEKI 366
QY 366 LPPAOLSVYTGILPDQVANSVGRLEWYLIRARFKMKELVPRNVERPEETRGALVLE 425
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 367 LPPAOLSVYTGILPDQVANSVGRLEWYLIRARFKMKELVPRNVERPEETRGALVLE 426
QY 426 PLGCVHENIAVDPSGMPINIKINVGPTLVPRGKSGEC--GCWEAPEYKRRPFRCP 484
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 427 PLGCVHENIAVDPSGMPINIKINVGPTLVPRGKSGEC--GCWEAPEYKRRPFRCP 485
QY 485 GFEKTVLERLLELRKVRRAEMKKYPPDSPEYRLIDEROKALKVLANASYGMGSGARVY 544
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 486 GFEKTVLERLLELRKVRRAEMKKYPPDSPEYRLIDEROKALKVLANASYGMGSGARVY 545
QY 545 CRECAKAVTAMGHHILRPTAINIARLGLKVIYGDTSLSFTYDPRKVNPFKIKIEELG 604
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 546 CRECAKAVTAMGHHILRPTAINIARLGLKVIYGDTSLSFTYDPRKVNPFKIKIEELG 605
QY 605 EIKLEEVYKRLPFTBAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTRVVEIYKTS 664
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 606 EIKLEEVYKRLPFTBAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTRVVEIYKTS 665
QY 665 EVNKAVEYRKVYKLEBEGKPIELVIMKTLISKRLSEETTEAPRVAAAKMISAGRYS 724
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 666 EVNKAVEYRKVYKLEBEGKPIELVIMKTLISKRLSEETTEAPRVAAAKMISAGRYS 725

```

Db 649 KVEDAIKVVYKTVIPDLRRYNEFIEDLIIMKTIDKNLDEYVTAHPVNAKKAAGYLVS 708

Qy 725 PGDIGYIVYVGGRIISORAMPYFMVNDPSQIDVTYYVDHDIIPALRIIGYGTTEKKL 784

Db 709 KGVIGYIVYVGGRIISORAMPYFLIKKNIIDVEYIDKQIIPVAMRIIEGFGVSSSL 768

Qy 785 KASATGOKTLPDFLAK 800

Db 769 K---TGVDILSPFK 781

RESULT 3

Q9P9N1 PRELIMINARY; PRT; 781 AA.

ID 09P9N1

AC 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, last sequence update)

DE 01-JUN-2002 (Tremblrel. 21, last annotation update)

DE DNA polymerase B3.

DE Sulfurisphaera ohwakuensis.

Archaea: Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae; Sulfurisphaera.

NCBI_TaxID=69656;

OK NCBI_TaxID=69656;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=TA-1;

RX MEDLINE=20450680; PubMed=10997874;

RA Irai T., Kurosawa N., Itoh Y.H., Kimura N., Horinchi T.;

RT "Sequence analysis of three family B DNA polymerases from the thermocacidophilic crenarchaeon Sulfurisphaera ohwakuensis."

RL DNA Res. 7:243-251(2000).

CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE + {DNA} (N).

CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

DR EMBL; AB032376; BAA93703.1; -.

DR HSSP; P56689; 1TGO.

DR InterPro; IPR002064; DNA_pol_B.

DR InterPro; IPR004578; Pol2.

DR Pfam; PF00136; DNA_pol_B; 1.

DR Pfam; PF03104; DNA_pol_B_exo; 1.

DR PRINTS; PR00106; DNAPOLB.

DR SMART; SM00486; POLBC; 1.

DR TIGRFAMs; TIGR00592; pol2; 1.

DR PROSITE; PS00116; DNA POLYMERASE B; 1.

KM DNA replication; DNA-Binding; DNA-directed DNA polymerase.

SO SEQUENCE 781 AA; 90384 MM; 982570EAB30C372C CRC64;

Query Match 53.6%; Score 2246; DB 1; Length 781;

Best Local Similarity 54.3%; Pred. No. 1,38-140;

Matches 432; Conservative 141; Mismatches 203; Indels 20; Gaps 10;

Qy 6 FTUDDSTVEYVGEPOVITWIGIAENGERVLIIDNSFRYPALALPGADPROVAQRIRAL 65

Db 5 PFILDSYDVVENKRVYIWIWIDKEGNNVLEKKFRYPALVDDSYNIDEIRKILKL 64

Qy 66 SRPSPITGVDDDKRYGRBRVLRIRTVLPEAVREYRELVKVDGVEDLADIRPAM 125

Db 65 SKPSPITSIDVEEKRTFGSVKALKIETVPAVYRVDRDAVKIKGVKSLVLEADIRYM 124

Qy 126 RYLIDHDLFPFTWYRVEAEPLENGKGFVVDVYLVKSRPELYGALAPTKLPDLRIALF 185

Db 125 RYSDINLKPFYVIEAEVEIKEN-NFRVKVYELK-KINTLYD-----KIPELKVIAPF 177

Qy 186 DIEYVSKSGSRPERDPYIVAAKTDDGDEVLFIAEGODKRPREFEYVYKRYDPDIIIV 245

Db 178 DIEYVNNKSGSNPRDPYIIIGVWTKESGK-QFLADKDDIRAIRREFINFGVQYVDPDIIIV 236

Qy 246 GYNNNNHFDMPYLRRARLIGIKLDVTRVGAEPSTSVGHVSVGRNLVNDLYDAEEMPE 305

Db 237 GYNNNNHFDMPYLRRARLIGIKLDVTRVGAEPSTSVGHVSVGRNLVNDLYDAEEMPE 296

Qy 306 IKISLEEVAVYLGVMKSESVIINMWEIPDYWDPKKRPLLOAYARDVATYGLAEKI 365

Db 297 VKVKTLENIADYIGLVLPKEXRTIVEMWDIPKMWDEKKRIDILKYNLDDAKSAVILGEVF 356

Qy 366 LPPAIQLSYVTGILPDDVGAMSVGFLEWTLIDPAEFKMKELVNNRVERPEETRGALVIE 425

Db 357 IPRGIELTRISGHPDQLSMASVGHREVMILMEAYKYNELINKEEREYESYEGGLVIS 416

Qy 426 PLRGVHENIIVLDSSYVPIMIKYVNGPDTLVPRPEKQEC-GCWEAPVGRFRCP 484

Db 417 PLRGVHEVYVLPSSYVPIMIKYVNGPDTLVK-----GECNCWSP-VGKFRKEPP 470

Qy 485 GFEKTVLERLLELRKRYRAEMKYPDPSPERYLLDEROKALKYLANASVGMGSGARWY 544

Db 471 GLYKIVLEKLIQERKQYKQKLMER-TIDEYDKRVLDARQRLKMANAIFYGMGLARWY 529

Qy 545 CRECAKAVTAMGSHLIRTAINTARKGLKVIYGDTSLFYTDPEKVENFIKIKELGF 604

Db 530 SKGAEAVTAMGSHLIRTAINTARKGLKVIYGDTSIFVKGGAD-INSILTEISSKFLG 588

Qy 605 EIKLEKRYKRLPFTFAKRYRAGLLEDGRIDIVGFPAVRGDMCELAKEVQKVEIYKTS 664

Db 589 EIKIDKTYKRVFTENKRYAGLLEDGKIDIVGFPAVRGDMCOLAQVNTVIELIKSG 648

Qy 665 EVKAVYVYKRYVLEEGKVPLEKLVIMKTLKSRLEEYTTAHPVNAKKAAGYLVS 724

Db 649 KVEDAIKVVYKTVIPDLRRYNEFIEDLIIMKTIDKNLDEYVTAHPVNAKKAAGYLVS 708

Qy 725 PGDIGYIVYVGGRIISORAMPYFMVNDPSQIDVTYYVDHDIIPALRIIGYGTTEKKL 784

Db 709 KGVIGYIVYVGGRIISORAMPYFLIKKNIIDVEYIDKQIIPVAMRIIEGFGVSSSL 768

Qy 785 KASATGOKTLPDFLAK 800

Db 769 K---TGVDILSPFK 781

RESULT 4

Q8ZVY2 PRELIMINARY; PRT; 785 AA.

ID 08ZVY2

AC 08ZVY2

DT 01-MAR-2002 (Tremblrel. 20, Created)

DT 01-MAR-2002 (Tremblrel. 20, last sequence update)

DE 01-JUN-2002 (Tremblrel. 21, last annotation update)

DE DNA-directed DNA polymerase (B3).

GN PAE2109.

OS Pyrobaculum aerophilum.

OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales; Pyrobaculum.

OX NCBI_TaxID=13773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=IM2 / ATCC 51768 / DSM 7523;

RX PubMed=11792869;

RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stettler K.O., Simon M.I.,

RA Miller J.H.;

RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum aerophilum."

RT Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).

RL EMBL; AE009857; AAL63952.1; -.

DR InterPro; IPR002064; DNA_pol_B.

DR InterPro; IPR004578; Pol2.

DR Pfam; PF00136; DNA_pol_B; 1.

DR Pfam; PF03104; DNA_pol_B_exo; 1.

DR PRINTS; PR00106; DNAPOLB.

DR SMART; SM00486; POLBC; 1.

DR TIGRFAMs; TIGR00592; pol2; 1.

DR PROSITE; PS00116; DNA POLYMERASE B; 1.

KM DNA-directed DNA polymerase; Complete proteome.

SO SEQUENCE 785 AA; 89547 MM; 8395EAI8C7804F1F CRC64;

Query Match 51.5%; Score 2156; DB 17; Length 785;

Best Local Similarity 53.7%; Pred. No. 1,28-134;

Matches 428; Conservative 111; Mismatches 210; Indels 28; Gaps 9;

Oy		9	LDSVYEVGKSPQIIVIMGIAENGEVTVLIDSPRPFFYALLARADAKVOAGRIIALSRP	68
Dd		8	LDAITSVVGCPVEIRITIGISSEGDNVVDNRFRPFYA-DCPCADBSSTRSQGLVA-P	65
Oy		69	KSLIIIGVEDDDRKRCGRPRPVLRITVLPEANREVELVANOVGYEDVLEADTRFAMRYL	122
Dd		66	VEEVAIVE---RRLTGPNFSFLIVAAVEEDVRKLDEAALAPGVSGVTEADTRFMRM	122
Oy		129	IDHDLPEPTTWRVAEPLEKNKGCFRVKDTLVKSREPPLYGA---LAFTKLPDIRILA	184
Dd		123	LDMGGVPPCSMTATVALETGBKLG-----NLIPYKAEMGWETGPFPPLRYLA	170
Oy		185	FDEIEYSKQGSPPRPDPVIVAIVKTDDGEVLFIAEGDKDRPIREFVEYVKRDYDPI	244
Dd		171	FDEIEYNRRGFPPDLARPVITLLAVASDGRVEFEVASGDRDSVRSIFDFVAEPDPVI	230
Oy		245	VGYNNNHFDMPYLRLRARILGILKLDVTNRVGAEPPTSVMGHVSVPGLANDLVDVAEMP	300
Dd		231	VGNNSQCDWMEVTLERARALGIPLYVD-RVIGAPOOSVYGHSVTSGRANDLXNIIDEFP	289
Oy		305	EIKIXSLEEVEVLEYGVWKSSERYINMMIEPDWDOPKKRPPLLQVARDVATYGLAEK	366
Dd		290	EIKLKLTLDRAVELEYGVWKKEEREVLPGHKIYEYWRDQGRPLRGVYIDDVYSTGLAEK	349
Oy		365	ILDPAIDISYVTGSLPDDOQGANSVGFLEWYLIRAFPMKKELVPRVPRPEEYRGATVL	422
Dd		350	LHPFLIQSSVSGPLDDOVAASVGNRIYEMMLIRAYRILGEVAPRREEREYFKCALIVL	405
Oy		425	EPLRGVHENIIVLDFESMTPNIMTKYNGPDLTVRSGEKSCGCCEMAPEVCHRPCCP	488
Dd		410	EPHPGIVSDVALDFSMPNPIMMKNYLSPTTYLBERGEDPFCGVYVAPRVEGRFRREP	465
Oy		485	GEPKYTLRELLELAKPRVAKEMKPSPDSPEVRLDROKALVLANSVGYMGMSAANY	544
Dd		470	GPIPLVQLIELKAKVREELKKYPBDSPEIRFLDERONALKIMANNMTGITGMGARMY	522
Oy		545	CBECKATANGRLIFRTAINIARKLGKTYIGDTSSLFYTDPEKVENFIKIYBELGF	604
Dd		530	KKEVASSTANARAILDVIEYARRAGLVYIGDTSSLFVKSGD-VEKLVKVEERKGI	588
Oy		605	EIKLEVKYKRLFTEAKGRVAGLEDGRIDIVGFEARNGDWCELAKEVOTKVEVLYLNTS	664
Dd		589	DIDIDXDSTYVLFTEAKGRVAGLEDGRIDIVGFEARNGDWSBLAKEVOLTVEILTISR	648
Oy		665	EVN---KAVEYAKTIKVELEGKVPIELKVINMKTLSKLELEYTTTEAPHVAAKMLTAG	720
Dd		649	DVSEAROKRVKRVAGVIDKRMYEVDLDDLIMKTLDKELDEYKAPPHHAILLKKRG	708
Oy		721	YRVSBDKIGIVYIKVGGGISORAMPYPMVWDQSQIDVVYUHQIIPALRIILFGFIT	780
Dd		709	YKKGATTIGIVVVKGEKESERRAVFIYFIDIEKIDLIDYYEQVIPALRIAYIGIK	768
Oy		781	EKKLKASATGOKTLFDF 797	
Dd		769	EGDLK-TGRSERITLDF 784	
<hr/>				
RESULT 5				
ID	Q9P9M2	PRELIMINARY;	PRT;	785 AA.
AC	Q9P9M2:	(TREMBLrel_15_Created)		
DT	01-OCT-2000	(TREMBLrel_15_Last sequence update)		
DT	01-JUN-2002	(TREMBLrel_21_Last annotation update)		
GN	Family B DNA polymerase.			
OS	Pyrobaculum islandicum.			
OC	Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;			
NCBI	TaxID=2277;			
FN	(1) SEQUENCE FROM N.A.			
RP	STRAIN=GEO 3;			

EX MEDLINE:02100754; PubMed:10633098;
R Kehler M, von Zastrow H-G.
RT Cloning and characterization of a Family B DNA polymerase from the
RT thermophilic Gram-negative bacterium *Picrobacculum islandicum*.
RL J. Bacteriol. 182:655-663(2000).
CC -1. CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
CC + {DNA} (N).
CC -- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR EMBL: AF195019; AAF27815.1; -.
DR HSSP; P56689; ITCO.
DR InterPro; IPRO02064; DNA_pol_B.
DR InterPro; IPRO04578; POI2-.
DR Pfam; PF00116; DNA_pol_B_1.
DR Pfam; PF03104; DNA_pol_B_exo_1.
DR PRINTS; PR00106; DNAPOLB-
DR SMART; SM00486; POLBC; 1.
DR TIGRFAMs; TIGR00592; poi2; 1.
DR TIGRFAMs; TIGR00592; poi2; 1.
DR PROSITE; PS00116; DNA_Polymerase_B; 1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase.
SEQUENCE 785 AA; 8948 MW; 10FB8B66A8F373D CRC64;

Query Match	Similarity	50.8%	Score 2129.5	DB 1	Length 785
Best Local Similarity	53.4%	Pred. No. 76-133			
Matches 429	Conserveative 141	Mismatches 195	Indels 39	Gaps 13	
Qy	9	LDSSEVYWGKEPOVITWGLAENGERVVLIDSPFPY---ALLAPGADPQVAAQIRAL	65		
Db	8	LDTTAVVAGSVPEIRIFGILSSSEERVLIDSPFPYVYDCAACEPA-----LKTAL	60		
Qy	66	SRKSPDIGVEDDKRYKFGSPRRVLIRITVLPEAVREYRELTKVNDVGEDVLEADIRFAM	125		
Db	61	SR-VAVIDDVQIVERRFLGSSKKFLKVLAKIPEDVRKLRAAAMSIPRVSGYEAIDIRFYM	119		
Qy	126	RYLIDHDEPFTTWRAEAPLEKNKGFVDVY--YLKSRPEFLVG--EALAPKLDLR	181		
Db	120	RYMIDGAVVPCSMNAVEE-----GRLGIGIPYVVSQ-----WYDIGECPSPS-----LK	166		
Qy	182	ILAFDEYVSKQSPREPRDPVIAVATDGDGDLVLAEGKDKRKLREFEYVSKYDP	241		
Db	167	VMAFDIEVYNNERSGPDPIRDPVMLAITNDGHEVEAAGKDDGVRARVDPDIRSDYD	226		
Qy	242	DIIVGNNHHDPMYLRARILGIKLVDTYRNGAETTSYGHNSVSGRLNDLYHAE	301		
Db	227	DVIVGNSNGDMYLRERKAVGCPKVD-RLSNPQOSVGHMSYGRANVDLYIVE	285		
Qy	302	EMPEIKISLEVAEYLGVKKKSERVIINMWEIPDYNDPKKRPILLOYADVRYATYGL	361		
Db*	286	EPPEIKLTLRVAEYGVKKREERLLPGKRIEYVQKPNKKRPILKIVLIDVRSITGL	345		
Qy	362	AERILPEAIOISYVTGLPLDQVAMSVGFRLEWYLIPAAFKMEILVPRNVRPEETRYGA	421		
Db	346	ADTLPEFLIOISSVGLPLDQVAAASVGNRVENMILLYARLGEVAPRRESEYEPKGA	405		
Qy	422	IVLEPLRGVHENIIVLDFSSMPYINIKYVGPOTLVRPEKSGEGECMAPEYKHFRR	481		
Db	406	IVLEPRPGVIEDVVLDFSSMPYINIMKKYNISPTLYLEPGEPPREGVNVAPVYGHFRR	465		
Qy	482	CPGPFKTVLERLLEKRVRAEMKKYPPSPPEKRLIDEBQKALKYLANASGYMGSGA	541		
Db	466	SPGFPVQLKSLVELKAVREAKKYPSPDEPKLIDEBQRLAKYMANNAIYGLMGVGA	525		
Qy	542	RMWCECKAKYANGRLILRTAINARLGLKYGTDSLPTTYDPEKVENFIIKIKEE	601		
Db	526	RMWKEVASVSTYARARILKDVTEQRRRIGIVVYGTDSLFPKKHG--VDLTKIYVEEK	584		
Qy	602	LGPEIKLEVKYKRLPFEAKKRYAGLLBDDGRIDVGFEAVRGMCELAKEVQTVAEIVL	661		
Db	585	YGDIKVKDXYAKLPTFAKKRYAGLLBDDGRIDVGFEEVVRGWSMLAKDVLQVLEIIVL	644		
Qy	662	KTSEVWKA-----VYVYKTVKVELSEGKAPIKLIVMTLSSKRLSEYTLAEAPHVVAAGRL	717		
Db	645	KASDIYEAHGVIKYIREILIRKTKYKKNINDILLMTLDEKLEDEKXVPPHYAAOITLK	704		

QY 718 SAGYRSPGDKIGYIVYKGGISORAMPYPMVDPQSDIVTYVDHGIIPALRIIGYE 777
 DB 705 RHGRVKGTTIGYIVYKGGKSERBALPYLLDDIKIDIDYIENQIIPALRIAEVT 764
 QY 778 GITEKJLKASATG--QKTLFDFLA 799
 DB 765 GKESDLK--TGMERSLIDFLS 785

RESULT 6

Q9P9K4 PRELIMINARY; PRT; 775 AA.
 ID Q9P9K4
 AC Q9P9K4
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN DNA POL.
 OS Pyrococcus glycovorans.
 Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 Pyrococcus.
 NCBI_TaxID=74610;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=AL585;
 RA Querrelou J., Cambon M.A., Lesongeur F., Portier P., Barbier G.;
 RT "Thermococcales genes organisation of species belonging to
 CC thermococcales and phylogenetic implications".
 CL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE
 + {DNA} (N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 DR EMBL: AJ004834; CAB81809.1; --
 DR HSSP: P56689; ITGO.
 DR InterPro: IPR002064; DNA_pol_B.
 DR InterPro: IPR004578; Pol2.
 DR Pfam: PF00136; DNA_pol_B; 1.
 DR Pfam: PF03104; DNA_pol_B_exo; 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR SMART: SM00486; POLBc; 1.
 DR TIGRFAMs: TIGR00592; pol2; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN 1.
 KM DNA replication; DNA-binding; DNA-directed DNA polymerase;
 Nucleic acid transferase; Transferase.
 SQ SEQUENCE 775 AA; 90492 MW; 677264920D770C0F CRC64;

Query Match 30.3%; Score 1268; DB 1; Length 775;
 Best Local Similarity 36.5%; Pred. No. 1; 1e-75;
 Matches 304; Conservative 154; Mismatches 278; Indels 98; Gaps 23;

QY 8 VLDSSEYVNGEPOVIIMGIAENGRVYLIDRSPPRYALLARQADPKQV---AQRIR 63
 DB 2 ILDDVYTEDKQPIIRIKP-KENGEFKYERNPFPYIYALKDQSDIEVKYKITAERHG 60
 QY 64 ALSRPKSPFIIGVEDDKRYKFRPRRLIRITVLPFAVEYRELYKNGVGVDFLEADIR 123
 DB 61 KIIV-----IYVEKAKKKKGLRPIEVKWLIEHNPQDVAIIDKIREHAAVVDIEFYDP 116
 QY 124 AMRYLIDHDLFPFTWYRVEAPELENKMGFRVYKVLVKSREPLVGEALAPTKLPDLIRL 183
 DB 117 AKRYLIDKGLI-----PMHG-----ELKTL 138
 QY 184 AFDIEVYKSGSPRERDPVIVIAKTDGDEVL-----FLAEGDDKRPREFREY 235
 DB 139 AFDIETLVEEB-EPAKPIIMISYADEGAKVITWKVDDPYVEVVSSEEMIKRPIKV 197
 QY 236 VKRYPDIIIVGNNHFPMPYLRRARILGIKLDVTRVGAEPPTSVGH---VSPGR 292
 DB 198 IREKDPDIIIVNGDSFPLPIVYKAEKLGKILPGRD-GSEPKKQRIKDMTAVIEIKRI 256
 QY 293 NVLDYVAEMPEIKIKLEVEAYLVGWKKSERYINMWEIPDYMDPKRPLLOVAR 352
 DB 257 HFDLVHVARITINLPYTLAEVYEAI-FGKPEKRYAH-EIAEMETGKGLERVAKYSM 313

QY 353 DDVATYGLAEKILPFAIOLSVYTGILPDQYGMVSGFRLEWYLIRAPFMKELVNRVE 412
 DB 314 EDKAVTYELGREFPWEADQSLRWGQPLMDVSRSSSGNVLKKAERNLANKPD 373
 QY 413 RPE-----ETVRGATVLEPLRGVHENIAVLDPSSWYPMIMKYKNGPDTLVAPGCKSE 466
 DB 374 EREYERLRRESVAGYVKEPEKLMGVLSDFRSLYPSIIITHNVSPTLNR-----427
 QY 467 CGCME---APEVYGRFRRCPPGFKTVLEKLELRVRVRAEMKKYPPDSEYRLDEROK 523
 DB 428 -GCMEYVAPAEVAGKFCCKPFGFIPSLKRLDEROBIKRMKA-SKDPIEKMLDYRQR 485
 QY 524 ALTVLNASVGMNGSGARVCEKCAKATYAMRHLIR-TAINIARKLGKVIYGGTDSL 562
 DB 486 AITILANSYGYGYAKARVYCKEACSVAMREYIEFVRKELEKFGKVIYDTDGI 545
 QY 583 FVTY---DPEKVB---NFIKIKEIG--FEIKLEKVVYRLPFTBAKKRYAGLLEDGRI 633
 DB 546 YATIPGAKPEIRKALKEFYETINAKLPGLLEFYEGFYRGFFV-TKKYALIDEGKI 604
 QY 634 DIVFEAVRGDWCELAKEVQTKVEIIVLKTSEVYKAVEYRKIVKLEBEGVPIEKLV 693
 DB 605 ITRGLIETVRDMSEIAETQAKVLEAILKRGVBEAVKIVKEVTEKLSKYEIPKLV 664
 QY 694 KTLKSLRETTAPRVYNAKRLMSAGYRSPDKIGYIVYKGGKISORAMPYPMV-- 751
 DB 665 EQITRPLHEKALGPHVAVKRLAARGVVRPQMVIGYIYLRDGPISKA---IIAEP 721
 QY 752 DP--SQIDVYVVDHGIIPALRIIGFYITEKKLKASATGQKTLDFPLAKSK 803
 DB 722 DPRKHKDADYIYENQVLPVAILAEFGYRKEDLMQTKQIGLAINLVK 775

RESULT 7

Q9HH06 PRELIMINARY; PRT; 775 AA.
 ID Q9HH06
 AC Q9HH06
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE DNA polymerase.
 GN POL.
 OS Pyrococcus glycovorans.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OX Pyrococcus.
 RN NCBI_TaxID=74610;
 RP SEQUENCE FROM N.A.
 RC STRAIN=AL646;
 RA Querrelou J.J.E., Cambon M.A., Lesongeur F., Barbier G.;
 RT "Thermococcales taxonomy and phylogeny based on the comparative use of
 RT 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase
 genes".
 CL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
 + {DNA} (N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 DR EMBL: AJ250335; CAC12849.1; --
 DR HSSP: P56689; ITGO.
 DR InterPro: IPR002064; DNA_pol_B.
 DR InterPro: IPR004578; Pol2.
 DR Pfam: PF00136; DNA_pol_B; 1.
 DR Pfam: PF03104; DNA_pol_B_exo; 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR SMART: SM00486; POLBc; 1.
 DR TIGRFAMs: TIGR00592; pol2; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN 1.
 KM DNA replication; DNA-binding; DNA-directed DNA polymerase.
 SQ SEQUENCE 775 AA; 90417 MW; 633F8392EDBEEC73 CRC64;

Query Match 30.2%; Score 1264; DB 1; Length 775;
 Best Local Similarity 36.5%; Pred. No. 2e-75;

Matches 304; Conservative 154; Mismatches 278; Indels 98; Gaps 23;

```

Oy 8 VSSSEVAVGPOVUWIGIAENGSRVULIDRSFPEYFALLAEPADPKOV---AQRIR 63
Db 2 LIDADITIEDOKPIIRIRFEK-KENGEFKEVIEDNRNRPYIALLKDDSOQIDVKKITNERBG 60
Oy 64 ALSRPKSPICVEDDRKRTKFGRRPRVRLIRFTVDEAVREYRELKAVNDGVEDLEADIRF 123
Db 61 KIVR-----IVVEKVKKKFEGSPRIEWMVLFEHPOVAPRIARIKIRREHPAVDIFPEVDIIF 116
Oy 124 AKRVYLIDHDLPEFTWREVAEPLENKMGFRVDKRYLVKSRPEPIYGAALAFTKLPDULIL 183
Db 117 AKRVYLIDGILL-----PMEGD-----BELKLL 136
Oy 184 APDIEVYSKQSGSPRPERDPYIAVAKTDGDEVL-----FLAEKGRDKRIRBPFEY 235
Db 139 APFIETLHEGE-EEFAGKPIIMISVADDEGAKYITKKKVDLPYEVAVSSREMIKRFPLK 197
Oy 236 VKRYPEDDIIYVYNNHNDMPYLLRPARLIGLKUDVRRVGAPEPTTSVGH--VSAPRL 292
Db 198 IREKQPDVYIITYNDSFDLPYLIVKAEKLGKPLGRD-GSPFKMQGLGMHTAVELKRI 256
Oy 293 NVDLDVYAEEMPEFKIKSLBEVAYELGYMKSSSRVLIIMMELPYMDPDKRRLLOQAR 355
Db 257 HDPLHVIRIRINLPYITLBNVYEL-FOKPKERYAH--ELAEEMETGGLERVAKISM 313
Oy 353 DDVRYATVGLAKLIRPLIOGSVYVGLPIDOYGAMSGFLEWYLIIRAFPMKELYPRAYE 412
Db 314 EDKATYELGEEFPEMAQDSKLVGQPLMVDSSSTGULVETIRKQYERLEAKNKPD 373
Oy 413 RRP-----FTYRBAIYELRPLAGVHENIANDLSSWYPMIKYINNVGDTLVARPBGKGE 466
Db 374 EREYERIRRESYAGGYKKEBGEVGLSADRSLSYPSIIITHNVSPDTLANE----- 427
Oy 467 CGGWE---APVYKGRFRPCPGFETVLTLELLELRKRVRAEMKKYPDSEYERLLDEROK 523
Db 428 -GREDVAVPRVKRKFCKOPFPISLTKRLDEROERIKRQKA-SKDPLEKKMDYBQR 485
Oy 524 ALKVLANASVGYMWSGAMRYCRECAKAVTANGRHILIR-TAINIARKLGILKYIGDPTSL 582
Db 486 AKITLANSYGYVYAKARWYCKEACBSVTAMGREYIEFVRKELEBEKFGFKVLYIDTGL 545
Oy 583 FVTV---DPEKVE---NFIKITIKELD--FEIKLEKVKYRLFETAKKKYVAGLEDDRI 633
Db 546 YATIGCAPDEIKKALFEFVYINAKLPGLEEYEGFVVRGFPV-TKKYVALIDEBEKI 604
Oy 634 DIVEGEAVRGMCCLAEVOTKVAEIVLKTSEVNKAVEVVRKYVELEBEKQVDEKLVIW 693
Db 605 ITRGLEIRDRMSIAKEQAVLELAILKNGVBAVKIVAEVTEKLSKTEIDPEKLVYI 666
Oy 694 KTLRKLEBEYTEAPHVAAKMSLSAGRVSPDDKIGYVYGGRLISGAPMYPMV-- 751
Db 665 EQGTRPLMEYKALGPHVAAKRLARNGKVRFGVNIIGYIVLKGDPILSKAA--ILAEF 721
Oy 752 DP--SODIVYVYNDHOLIPALRILIGYIGTEKKLAKASTGQKTLDFPLAKSK 803
Db 722 DPKRKYVAYEYIENVLPRVLILBAGIRKGLRMOKTKOTGLTAMLVKKK 775

RESULT 8
Q8MTW5
Q8MTW5 PRELIMINARY; PRT; 830 AA.
ID O8MTW5;
DT 01-JUN-2002 (TREMBLrel, 21, Created)
DT 01-JUN-2002 (TREMBLrel, 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel, 21, Last annotation update)
DE B family DNA polymerase.
GN POLB OR MK1039.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
NN [1]

```

RP SEQUENCE FROM N.A.
PC STRAIN=AV19 / DSM 6324 / JCM 9639;
EX MEDLINE=21927647; PubMed=11930014;
RA Steerey A.I., Mezneva K.V.V., Makarova K.S., Polunin N.,
Rastserbina O.V., Shakhova V.V., Belova G.I., Afanudin L.,
Rakate D.A., Kozozin I.B., Tatuzov S.I., Wolf Y.I., Stetter K.O.,
Malykh A.G., Koonin E.V., Kozayvkin S.A.;
RT "The complete genome of hyperthermophilic Methanopyrus kandleri AV19
and monophyly of archaeal methanogens";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).
DR EMBL; AE010392; AAM02522.1; ...
KW Complete proteome.
QO SEQUENCE 830 AA; 94561 MW; 094B15D7246EE47 CRC64;

```

QY 19 EPOVILWIGIAENGERVVLIDRSFREPYPALLAPGADKOYAOKIIRALSRPKSPITIGVED 78
D 14 EBPVILVIRREDGDPALVVKGRGRRPYAAVEDEFPDPSBV--EKL-----SGVEVEEV 65
QY 79 --KRYKFRPRRVLRIKTVLPEAVREYELVKNVDGVEDLLEADIRFAMRLIJDHDFPP 136
D 66 LLEHPYGDGRDREVLIRIVATYPKVVPKLEQVCKLIDGVEKEVEYENDIPFRRAAVDLNLPFA 125
QY 137 TWYRPEADPLN-----KMGFRVDKYVLKSRPEPLVGEALAPFKPDLIRIAPDIEV 189
D 126 S--EVDVSDLDLTGSMGSLPAYFADVEDRARELDHFPY-----IEDLVASFPLEV 173
QY 190 YSKOSP--RPRBDVVI--AVYTDGDEVEFLIA--EGKD-----DRK 227
D 174 LAEPGTITKASGPIIISPAVSTPDGRRANVYITWKGDSEBSEVDEVEIVCRSEAA 233
QY 228 PIREVEVVKRYPDIIIVGNNHNDPMWYLLRRLILGKLD--YRRRGAPEPTSYNG-- 284
D 234 ALRRFFDEFRVYDDVFTYNGDEFDLYLQHRGKGLIDVSLRPAKGGIILKKG 293
QY 285 --HVSPPRLAVDLYDAEBMPEIKISLE--VAEYLGVMKKSERYIINMEIDPYDDP 341
D 294 RIASDIFRAHVDLYHARKNLKERFPLEBAVDVAVGEKEMELA--DINEAA-- 346
QY 342 KKRPL--LLOYARDVATYGLAEKILFPAQLSIVYTGLEPLDOVGAMSVGLEMYLIRA 399
D 347 KGNIDELMRYSADAHYTLLELELAOVLELSYLRPLPLPATRFRFGQJLAEM--RA 403
QY 400 AFRMKE--LVNNVERPE-----ETYRGALVLEPLRGVHENIIVLFPSSMYENIMIK 450
D 404 IYKAOEDILVNNKPTDEYKRRRRKAYKGAIVPEPEIGLHEVVCVDPALYENVAH 463
QY 451 NYGDPVLVRPEKCGECGCWE-----APEYKFRPRCPGFSFKTVLRL 495
D 464 NISDPTF-----DCCCPVYVEVDPTDAVPAVDVGHKFCRRKRGFFPLVEGL 515
QY 496 ELKKVRAEMCKYPPS--PEYRLLDEQOKLYLANASGYMGWSGARWCRCSCAAVY 553
D 516 ERRRLKRRKLKLDTESHRHEAKIIDVQOAYKVLNASTYGYMGANARWCRCSCASV 575
QY 554 ANGBHLITANIN--RKLGLKTYGDTSLVYTD-----EYVNFILYIEELGFE 605
D 576 AMKRYIYSEVRLIEEKIGLKYVYDGTOSLVKPLPDADLEETIERKEBLEVNGRPLVE 635
QY 606 IKLEBYKRLFTFEAKKRYAGLLEDGRIDIVGFEAVNGDMCEBLAEVOTKVEVILVLTSE 665
D 636 LELEDAVYRILIV--TKKRYAGTDEGKLVYTGLEVARMDAPARLETORRLKILDND 694
QY 666 VNKAVEYVRKIVKELEBGVPIEKLVIWKTLSKRLEBYTTEAPHVAAKRL--SAGRYVS 724
D 695 PEAALKEIHEVLEBKSGDVDDIDELAVTSQLTKKPSEYVOKSPHRAALRIARHLGVEPE 754
QY 725 PCDKIGYIVVGGGRISORAMEYFWKDP--SQIDVTVYVDHOIIPALARIIGVYGLEKK 783
D 755 PGIVIRYIVVGPSSVDAKAYVLELVEEGEPVPVDYIEHOLIPAVRIRALIGYSRGO 814

```

QY 784 LKASATGOKTLFDF 797
 DB 815 IVGETASOKTIDQF 828

RESULT 9

Q9HH98 PRELIMINARY; PRT; 759 AA.

AC 09HH98
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA polymerase (Fragment).
 GN POL.

OS Pyrococcus sp. (strain ST700).
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.

OX NCBI_Taxid=69015;

[1]
 SEQUENCE FROM N.A.

STRAIN=ST700.

RA Querrelou J.J.E., Cambon M.A., Lesongeur F., Barbier G.;
 RT "Thermococcales taxonomy and phylogeny based on the comparative use of
 RT 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase
 RT genes.";

CC Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
 + (DNA) (N).

CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

DR EMBL; AJ250332; CAC12847.1; -.

DR HSSP; P56689; IRGO.

DR InterPro; IPR002064; DNA_pol_B.

DR InterPro; IPR004578; GPCR_Rhodopsn.

DR Pfam; PF00136; DNA_pol_B; 1.

DR Pfam; PF03104; DNA_pol_B; 1.

DR PRINTS; PRO0106; DNAPOLB.

DR SMART; SM00486; POLBc; 1.

DR TIGRfam; TIGR00592; Pol2; 1.

DR PROSITE; PS00116; DNA_POLYMERASE_B; UNKNOWN_1.

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

KM DNA replication; DNA-binding; DNA-directed DNA polymerase.

FT NON TER 759

SEQUENCE 759 AA; 88431 MM; A8459B6AAD8F3A3 CRC64;

Query Match 28.8%; Score 1208; DB 1; Length 759;

Best Local Similarity 35.4%; Pred. No. 1e-71;
 Matches 289; Conservative 156; Mismatches 273; Indels 98; Gaps 21;

QY 353 DVVATYGLAKLIPFAIOHSYVTLGLPLDOVGAMSVGRLEWYLIRAFKMLVPRVRE 412
 DB 314 EDKAVTYELKEFFPMEAOQLARLVGGPMDVSSSTGNLWVFLRRKAYRNEELAPKPD 373
 QY 413 RPE-----ETRGATVLEPLRGVHENIAVLDFSSMYPNIMIKVNPDPDLVRGKCGE 466
 DB 374 EKEVEKRLRESYEGGYKEKEKGLMEGIVSLDFRSLPSLIIITHNVSPTLNRG----- 428
 QY 467 CGCM-EAPVKAHRRRCPPGFPTVLERLLEKRRVAEMKKYPPSPPEYLLDERQAL 525
 DB 429 CGKDEAPEVGHFCKDPFGFIPSLGLDLLEBKOKIKRMEK-SKDPLEKKLLDYQRAI 487
 QY 526 KYLANASGYMGSGAMWYGRCAKATANGRLIRAINIARKLGKVIYGTDSLFT 585
 DB 488 KILANSFYGYGAKAWYCEKCAESVTANGROYIELVRLEBERGRKVLITDGLYAT 547
 QY 586 YDPEKVENFIKIKELGF-----EIKLEKYYRLPFTFAKKRYAGLLEDGRI 633
 DB 548 IPGEK--NMEIRKALKEFVNYINSKLPGLLELEYBGFYTRGFV-TKKYALIDEGKI 604
 QY 634 DIVGFEAVRDWCELAKEVGTQVVEIVLKTSEVNAVEYRKIVKELEBGKVPDIKLVIM 693
 DB 605 ITRLEIVRDMSEIAKETQKVLKILKGNVEAVKIVKVEYTKLSNVEIPKLVYI 664
 QY 694 KTLSKLREYTTTAPHVAAKRMLSAGYRSPDQIGYIVYKGGGRISQAMPYVVKD- 752
 DB 665 EQLTRPINEYKALGPHVAVKRLAARKIKIRPGMVGIVGLLRDGPISKBA---IAIEEF 721
 QY 753 ---PSQIDVTVYVDHQTIRPALRILIGFGITEKKL 785
 DB 722 DGKHKTDAYEYIENQVLPAVERILKAFGYKKEDLR 757

RESULT 10

Q8TSB3 PRELIMINARY; PRT; 937 AA.

AC 08TSB3;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA-directed DNA polymerase.
 GN MA0885.

OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.

OX NCBI_Taxid=2214;
 RN [1]
 SEQUENCE FROM N.A.

RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=1932238;

RA Galagan J.E., Nussbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Ancoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., DeArjano J., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Camm I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity.";
 RL Genome Res. 12:532-542 (2002).
 DR EMBL; AB010753; AAM04324.1; -.

KM DNA-directed DNA polymerase; Complete proteome.

SEQUENCE 937 AA; 105600 MM; 4ADAB7F59AD59B6 CRC64;

Query Match 20.4%; Score 852.5; DB 17; Length 937;

Best Local Similarity 31.4%; Pred. No. 5.7e-48;
 Matches 294; Conservative 132; Mismatches 324; Indels 185; Gaps 43;

QY 293 NVLDYVAEMPEIKIKSLAEVAEYLVGMKKSERYINMWEIPDVMDPKPKRLLQYAR 352
 DB 257 HFDFLPVIRRTINLPYTLRTVEAI-FGKPKKVPYH-EIAEMETGKGLERVAKYSM 313

```

QY 6 FTVLDSSVEYVKE--POVITWIGIAENGERVVLIDRSFPFYALLAAGADPKQVQNRIR- 63
DB 5 FQILDADVEYVNDSCGVIRIRLFGAGADKSCVCCFPVDFEPYTY--LKASGDIAVAVRLIKD 62
QY 64 ALSRKSPITIGVEDDKRYKGRPRARLRTYLPANAREK--ELVK-----NVDGVEDVL 117
DB 63 TFEQYKKEIVEKEFEPYQKTKCKMLRTYTLPRQVPELIDELIKRDVLAAGQWQY 122
QY 118 EADIRFAMRYLIDHDLFFPFYVTRVEAPLE--NKGK-----FRVDKYVYK--S 162
DB 123 ESDILFNRFLIDRALGQWVASBESKPPDPVRYIAGASAMRSCENFACDSAVLASGLK 182
QY 163 REEPLYGHALAFTKLPDLRIALADIEVYSKQS--PRPERDVIYIAVKTD----- 211
DB 183 RVENL--ALAP-----LKYLAFLIECLPLDGMSPDPVSPIIMISFSEPEYKQHKTLI 234
QY 212 -----DGEVLFLAEGKDKRPIREFVEYVYKYPDIIYGVYNNNHEDWPFYLRBAR 262
DB 235 ILAKEPAAGMDGD-----VLSCMDETEMLNKFEIICEYDPDIYAGYNNHQPDIPIYTERVK 290
QY 263 ILGILKLDVTRVGAEPITTSYHGVSVP-----GRLANDLYDAEEMPEIK 307
DB 291 ALVAK-----GETINSVVGSDSPICYRKFGILITREMKGRVYVDALPLVRAAFGLK 342
QY 308 IKSLEEVAYGVWKKSERIIMMIEIPYMDP--KKRPILLQYADDPVATYGLA--EK 364
DB 343 QYTLRAVSKEL-----LSREKLDVPRLEMEHNDSGDKRKVDYARQDSELALEVLRLR 399
QY 365 IIPFAIQLSYTYGLPLDQV--GANSVGRLEWYILRAAFKMKELVNRVERPE----- 415
DB 400 LLDKTYALAOVSGSLQELIYDGGQTSW--VETLLR--EQLKD-----RVLLPKQDELIS 451
QY 416 -ETY-----RGATVLEPLRGVHENIAYLDFSSYPMIMIKYVNPDTLV--RPEKCC- 464
DB 452 AERYDMSDLKGGVLEPKKGLLENVLLIDYKSLPYITMAHMLCTYTVTRDRPQKTI 511
QY 465 -----GEGCWAEAEVGRFRRCPPGFYTVLERLLRLKRAVRAEMKKYPPSPSEYRLD 519
DB 512 KPPSGE--FVPEV--FR-----GIYPSIIEIDLNNKRGDTKRMKR--TSDENEHRYLD 560
QY 520 ERQKALKYLANASYGWMGSGARWYCRECAKAVTAMGR--HLIRTAIYARKLG----- 571
DB 561 ATQALIKILINSFYGYGARALYSLTLANAVTSFGRSNLTNRDLINGRIGKIVLANS 620
QY 572 -----LKVYGPDTSLFY-----TYDPEKV--ENPIKIIXEL 602
DB 621 AALLLEAGKLSQDRIVELSVAYGDTDSVFHCAKQDLSLEVSIVGRSLSEIYASL 680
QY 603 --GFEIKLEKYKELFTTEAKKRYA--GLE-----DGRIDYGFAYVGNDCMLAKEXQ 653
DB 681 PDPMELEFESVAKRALLI--AKKRYALMLFEPNNSGEMKIKVGMETVARDWCBLISITL 739
QY 654 TKVVEIYVUKTSEVKKAVEYRKI--VKELEBKYP--TEKLYIMKTLSKRLEEYTEAP 708
DB 740 NNVLEFVLIEGDDVKAHEVRKAVSVRNLDPGKAGIIEELVITRLTRKADSKYKNQP 799
QY 709 HVAAKRMLS--AGRYVSPGDKIGYIVIKGGGRIISGRAN--PYFWVKPSPQIDYTYVDQI 766
DB 800 HLTVAENLKRRKTGIMPISIGRIRPFVITACKGLFVDRADDPVYRENNVPIDYDYVYKQI 859
QY 767 IPAALRIYFGITTEKKLKASATGQKTLFDFLAKK 801
DB 860 LPVVERILEVGVVMSLSLDPDAK--QKGLDFEVKK 893

```

RESULT 11

ID Q96H98

PRELIMINARY;

PRT: 1107 AA.

AC Q96H98:

01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001

(Tremblrel. 19, Last sequence update)

DE Polymease (DNA directed), delta 1, catalytic subunit

```

DB (125KD)
OS Homo sapiens (human)
CC Bakayofa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_Taxid=9606;
RN (1)
RP SEQUENCE FROM N. A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N DIOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
+ (DNA) (N).
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
CC ALPHA, BETA, GAMMA, DELTA, AND Epsilon WHICH ARE RESPONSIBLE FOR
CC DIFFERENT REACTIONS OF DNA SYNTHESIS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR EMBL: BC008800; AA008800.1; -.
DR InterPro: IPR002064; DNA.pol.B.
DR pfam: PF00136; DNA.pol.B.1.
DR TIGRfam: TIGR00592; pol2.1.
DR PROSITE: PS00116; DNA POLYMERASE B: UNKNOWN 1.
DR DNA replication; DNA-binding; DNA-directed DNA polymerase.
KY SEQUENCE 1107 AA; 123611 MW; 10A9A3271916EDBA CAC64;
SQ
Query Match 18.2%; Score 760.5; DB 4; Length 1107;
Best Local Similarity 28.7%; Pred. No. 9.4e-42;
Matches 255; Conservative 144; Mismatches 320; Indels 169; Gaps 33;
QY 17 GKSPQVITWIGIAENGERVVLIDRSFPFYALLAAGADPKQVQNRIR---ALSRK--- 69
DB 120 GSVPLVIRAGVYDSEFVSCVCHIGFAPYFTYAPGPFSEHMGDLQREMLAISPSRSG 179
QY 70 -----SPIIGVEDDKRK-----YFGR--PRVLRIRTVLPEAVREYELVKNVDVEDV- 116
DB 180 RELTGAVALVAVELCSRESMEGYGHGSPFLRITVALPRLVA PARRLLEGIRVAGLGT 239
QY 117 -----LEADIRFAMRYLIDHDLFFPFYVTRVEAPLENNKQGFVVDVYV- 164
DB 240 SPAPYAVANDFEIRMTVDIIVGNNMELPARKVALRLKEKATGCGLEADVLSVDVSHR 299
QY 165 EPLYGALAPTLPLDRIALADIEVYSKQS--SPREPRPVI--AVKDDDEVLFYA 220
DB 300 PBPMPKRIIP-----LKVLSFDIFCAKRGKIPPEERDPVQISLRLKRGEPFLRLA 354
QY 221 -----EKQDKRPTREREYEVYKYPDIIYGVYNNNHEDWPFYLRRA-- 261
DB 355 LTRPCAPILKAVQSYKEBDL--LQAWSTIRIMDDPVITGYNIGFPLPYLISRAQT 412
QY 262 -----RLIGIKLDV-----TRVGAEPITTSYHGVSVGRQLNVDLYAEEMPE 305
DB 413 LKQTFPFLGRVAGLCSINRDSFOSKQGRDITV--SVNNGRVQMDMLQVL--LRE 466
QY 306 IKIKS--LEEVA--EYLGMKKSERYIINMMEIPDYMDPKRPLLYOARDVR--ATY 359
DB 467 YKLRSTYLANAVSFHLEGGKEDVQHSI--ITD-----LQNGRDQTRRLAVY 511
QY 360 GLAEKILRP-----AIGSYVTGLPLDQVANSVGRFLEWYILRAAFKMELEVP 408
DB 512 CLADAVLPRLLERLMVAVNAVEMARVGPVLSYLSGGQGVKVVSGLQNAHBEGLMP 571
QY 409 NRVREEREYRGATVLEELRGVHE--NIAVLDSNMYPMIMIKYVNGDPTLVRG--EKGG 465
DB 572 VVKSBEGBDYQATVIEPLKGYDVIALDLSSLSYPSIMAHMLCTTILIRGTQAKG 631
QY 466 EGCWAEAEVGRFRRCPPG--FFKT-----VLERLLLEKRYARAAEMKKYPPDSE 514
DB 632 -----LTBQFLRTPTGDEFPKTSVRKGLLPOLLENLSARKRAKAELEK--ETDPLR 682
QY 515 YRLDERQALKYLANASYGWMGSGARWYCRECAKAVTAMGRHLI-----RTAIN 565
DB 683 RQVLDRQALKYVANSVYGTGAQVQKPLCLEISQSVTGFGRQMIKTKQLVESKYTVE 742

```



```

QY 566 IARKGLKVIYGDPTDLSFVTYDPEKVENFIKIKEELGF-----EIKLEKYKRLF 616
DB 743 NGYTSKAVYIGDDTSVWCRGVSVAEMALGREAADVSGHFPSPRLFEKYPF-Y 801
QY 617 PTEAKRYAGLL-----EDGRIDIVGEFAVRGWCMLAKEVQFVKEIVLKTSEVNKAV 670
DB 802 LLISKRYAGLLFSSRPDHRMCKGLEAVRDNCPVLAVNLVTASLRLLIDRDPGAV 861
QY 671 EYAKRYVELEEGKVPLEKLYIWKLSKRLSEYTTAPDHVAAKM--LSAGYRVSPGDK 728
DB 862 AAADVDISDLGNRHDISQILVITKELTPAADYAGKQAHVELAEMRRKRDGSAFSLDR 921
QY 729 IGYIIV---KGGGRISQRAWPFYMWKDSQIDVYVYVHOIIPALRI 773
DB 922 VPYIISAKGVAAVMKSEDPFLVLEHSLPIDTQYTLLEQOLAKPLRLI 969

RESULT 12
PVLTO
PRELIMINARY; PRT; 1105 AA.
091VTO
01-DEC-2001 (Tremblrel, 19, Created)
01-DEC-2001 (Tremblrel, 19, Last sequence update)
01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE DNA polymerase delta 1, catalytic domain.
GN POLDI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
ON NCBI_Taxid=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST TUMOR;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
+ {DNA} (N).
CC -I- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR
DIFFERENT REACTIONS OF DNA SYNTHESIS (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
DR EMBL: BC009128; AA009128.1; -
DR MGI: 97741; PolDI
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR004578; Pol2.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
DR TIGRFAMs: TIGR00592; Pol2; 1.
PROSITE: PS00116; DNA POLYMERASE B; UNKNOWN 1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase.
SQ SEQUENCE 1105 AA; 123744 MW; 20323690DD472406 CRC64;

Query Match 17.8%; Score 743.5; DB 11; Length 1105;
Best Local Similarity 27.9%; Pred. No. 1.3e-40;
Matches 246; Conservative 153; Mismatches 319; Indels 165; Gaps 33;

QY 20 PVIITWGIANGERVVLLIDRSFRPYFVALLPAGADPKQVQRIIR---ALSRLP----- 69
DB 121 PILRAFGVTDGFSVCHIQGFAPFYTPAPPGFGEHLSEHQELNAISRDSQKREL 180
QY 70 --SPICGEDDKRK---YFGR-PRRVLRIRTVLPEAAREVRELVK---NVDGVD----- 115
DB 181 SGPAVLALELCSRESMFGHGHGSPPLRITLALPRILMAPARLLLEQGVRFGLGTPSFA 240
QY 116 VLEADIRFAMRYLLIDHDLPFTW-----YREAPELENKGFVVDKYVL-VKSRPPL 167
DB 241 PYEAVNDVEIRFMVDADIVGCNMWLELPAKGYVRAREKATLCQLEVDVLMGSHVSHPEBG 300
QY 168 YGEALAPKLPDLRIALVDIEVYSKQ-SRPBRDPVIV---AVKTDGGEVLFIA--- 220
DB 301 OMORIAP-----LRVLSFDIECAKGRKGFPEPERDPVIGICSLGRMGEPEPFLALTL 355

```

```

QY 221 -----EGKDKRKPIREFVSVRYDPDIIYGVNNNNHDPMYLARRA----- 261
DB 356 RPAAPILGAVQSYEREDL--LQAMADFTLAMPDVTGYNINQFDPYLSISAQALKV 413
QY 262 -----RIIGIKLDV-----TRVGAEPITSVGHGVSVPGRILNVLYDAEEMPEIKI 308
DB 414 DRFPPLRVTVGLNRINIDSSFGSRQVGRDSKV-----ISWVGKQVMQMLQVLRHKLRS 469
QY 309 KSLLEVA-EYLGVMKKSERVIINMMWIPDYWDPKRRPLLQYARDDVR--ATYGLAEK 364
DB 470 YTLNAVGFHFLGEOQEDVQHSI-----ITD-----LQNGNEQTRRRRLAVYCLKDA 514
QY 365 ILPPA-----IQLSVTVGLPLDYGANGSVGRLEMYLIRAFKKEVLPPNVER 413
DB 515 FLPLRLRLRLMLVNNVEMARVTVGPVPLGYLLTRGOQVAVSOLRQMRGGLMPVVKTE 574
QY 414 PEETRGAVILEPLRGVHE-NIAVLDFSSMYPNIMIKNVGPPTLVLPFG--EKQEGCGW 470
DB 575 GGEDYTGATVIEPLKGYVDVPIALTDFSSISYPSIMAHNLCYTTLLRPGAQKLG----- 629
QY 471 EAPF-----VHRFRRCPPGFEKTVLERLLELRVRRAEMKKYPPDSPEYRLD 519
DB 630 LKPDFFIKTPTGDEFVKSVRK---GLLPQILEVLSARRAKAEILAQ-ETDPLRRQVLD 685
QY 520 ERQALVILANASGVNMGARVYCECAKATYAMGRHLIRPAINARK-----L 570
DB 686 GRQALTKVANSVYGFYGAOVGKLPCLDISQSVGFGRQMIETKQOLVESKYTVENGDA 745
QY 571 GLKVIYGDPTDLSFVTYDPEKVENFIKIKEELGF-----EIKLEKYKRLFTEAK 621
DB 746 NAKVYIGDDTSVWCRGVSVAEMALGREAADVSGHFPSPRLFEKYPF-YLLISK 804
QY 622 KRYAGLLIEDGR-----IDIVGEFAVRGWCMLAKEVQFVKEIVLKTSEVNKAVERYK 675
DB 805 KRYAGLLFSSRSDAHDMDCKGLEAVRDNCPVLAVNLVTASLRLLIDRDPGAVAAKQ 864
QY 676 IYVLEEGKVPLEKLYIWKLSKRLSEYTTAPDHVAAKM--LSAGYRVSPGDKIGYI 773
DB 865 VISDLGNRHDISQILVITKELTPAADYAGKQAHVELAEMRRKRDGSAFSLDRGVYVI 924
QY 734 V---KGGGRISQRAWPFYMWKDSQIDVYVYVHOIIPALRI 773
DB 925 IGAAGVAAVMKSEDPFLVLEHSLPIDTQYTLLEQOLAKPLRLI 967

RESULT 13
08XON7
PRELIMINARY; PRT; 1104 AA.
08XON7
01-MAR-2002 (Tremblrel, 20, Created)
01-MAR-2002 (Tremblrel, 20, Last sequence update)
01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE Probable DNA-directed DNA polymerase III.
GN B10H4.020.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
ON NCBI_Taxid=5141;
RX [1]
RP SEQUENCE FROM N.A.
RC German Neurospora genome project;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RA EMBL: AL670010; CAD21389.1; -
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR004578; Pol2.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
DR PRINTS: PR00106; DNAPOB.

```

Wed Apr 23 08:09:19 2003

us-10-034-849-2.rspt

Page 10

DR SMART: SW00486; polbc: 1.
DR TIGRFAW; TIGR00592; pol2: 1.
DR PROSITE: P50016; DNA_POLYMERASE_B; UNKNOWN_1.
KW DNA-directed DNA polymerase.
SQ SEQUENCE 1104 AA; 12507 MW; 385A1A08FD26042F CRC64;

Query Match 17.6%; Score 737.5; DB 3; Length 1104;
Best Local Similarity 28.6%; Pred. No. 3.1e-40;
Matches 261; Conservative 154; Mismatches 311; Indels 185; Gaps 39;

17 GKEPOVYINGIAENGERVVLIDRSFRPFY--ALLAPGD--PKOVAORIALSRP---- 68
123 GGGATIKLFGVTENGNSVLLVHKDFKHYLVQAIVSFGDDCPKRAFLLETQLMPTPAI 182
69 -----KSPIGVEDDKKRYGRRRLVLRITVLPEAVREYRELKGV 110
183 HSYLTWRENNYGFQNTQNPYIKVTVNDPKFLPRVRLILETNKA-----NMKGWKS- 235
111 DGVEDVLEADIRFAMRYLIDHLPFTWYVEAE-----PLENKGFRVDKXY 158
236 DGSIMTFD-DIGYLRFWDCSIGMSVLEAPAGSYQVINDGHSNCOLEAMSYR-----D 290
159 LYSRPEPLXGEMALPTKLPDLRIAPDIEVSKG--SRPERDVIYA-VKTDGDEV 216
291 LIAMK---VGEW---SKMAPRLISFDIEGAKGIFPEQPHDSVIOIANIVTKYGEK 344
217 LFLAE--GKDDRKPI-----REFVEYKRYDPDIIIVGNNHFMWPLYLR 259
345 PPMANVFCLDTTSPVIAQIISFREDMDLREGNRLLEQVDDPIIGYNINAPDFPYLD 404
260 RARILGK-LDV-----TRVGAEPFTSVGHVSFGRNLVNDLYDAE 302
405 RAKILKVKDFDFMSRTVRVSVAKETNFSSKMGNRDTKATNN-----RLQLDMLQVQK 460
303 MEIKIKLSLEV-AEVLGVWKS--ERYINMWEIPDYDDPKRPLLYOARDVATY 359
461 DYGLRSYTLNSVCSHFLGKEDVHSHMTELE-----GTPESRRRLALYCLMDAVLP 515
360 GLAEKI--LPAIOLSVYTG.LPDOVGMSVGRFLMXYLIRAFKMKELVN--RYEPER 416
516 RIMDKLSCLENTEMARVATGVPFNFLARQOKPISQFLKLEOKIVINLRPSSSE 575
417 TYGATVLEPLAGVH-NIAYLDSSMTNIMIKVNGDITLVREGKCGEGCEWAPR- 474
576 QYEGATVLEPTKGYIDVPLATIDPASTLPSTIMOHNLCYTTLI-----KKRIDERWDLKD 631
475 -----VGRFRRCPPGFKTVLELRLLELRVAEMKKYPPDSPERYLLDE 520
632 EDYIVTNGDMFVTTKQ-----KGLLAQILEELLSARQARREL-AVERKDPFKKAVLNG 685
521 ROKALKVLNANASGYGWSGAWYCRECAAVTAMGRHLI-RTAINIARKLGL----- 572
686 ROLALKVANSVYGLTNGTLFCLLEIASVTSFGROMIERTKEVEERYTIANGSHD 745
573 -KVLGPTDLSLFYTDPEKV-----NPKIKIKELGEILKEVYGRLE 616
746 AQLVYGDTSVMWKGSTDLAEAMKLGODASEVYSKFIKPIK-----LEFEKVYPPYL 799
617 FTAKKRYAGL-----EDGRDIVGEAVRGWGLAEVQTKVLEVLKTSVYKAVE 671
800 LIN-KKRYAGLWTKREKDKMDTKGJETVARNCLLVQVLEKLRNMLIDQVPAQA 858
672 YVKLYKELEBEGVPLEKLVWKLTKRLEBYTEAPHYAAKRM--LSAGYRVSPDKT 729
859 YVKQTLADLLQNKIDMSKLVITKALTK--ENYDAKAHVELLQNRKKKADASAGLDV 916
730 GYVIVG--GRLSORAW-PYFMVYDPSQIDVTYVHOIIPALR-----ILGFGITKEK 782
917 AYVWVGATGSKNFERSEDPYVLEHNVPIDTKYVLDNOLAKPLRIFEPILG-----ET 971
783 KLASATGQKT 793
972 KASLSLHGHT 982

RESULT 14

08SQPS PRELIMINARY; PRT; 974 AA.
ID 08SQPS
AC 08SQPS; 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DNA polymerase delta catalytic (large) chain.
GN EC009_0430.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_Taxid=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA MEDLINE=21576510; PubMed=11719806;
RA Kadiuka M.D., Duprat S., Cornilliot E., Metenier G., Thomarat F.,
RA Prensier G., Barbe V., Peyreallade E., Brotier P., Winkler P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivas C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL; AL590451, CAD27015.1; -;
SQ SEQUENCE 974 AA; 111628 MW; 957E519E6EBB455 CRC64;

Query Match 17.5%; Score 732.5; DB 5; Length 974;
Best Local Similarity 28.0%; Pred. No. 5.6e-40;
Matches 249; Conservative 152; Mismatches 320; Indels 167; Gaps 33;

2 TEVYFVLDSS--SYEVNGKEPOVYINGIAENGERVVLIDRSFRPFVALLAPG-----AD 54
10 THIFVNDSDGLSGSC---PVFSIFGNYSKGPVAVLKNPFVYVRSNGKYEKEED 65
55 PKQVAGIRALSRPSPPIIGVEDDKRK---YGRPRVLRITVLPEAVREYRELKGV 110
66 IKESVORLDV---KATILEVAVMKOSILGYEGKTRVYRLTNTPHVSTALKVLLSEG 121
111 DGVED-----VLEADIRFAMRYLIDHLPFTWYVEAEPLNKGFRVDKXYLVKSRP 164
122 ISVGEKVRPRVYESNPFVLRFCMDLGIWMSYLVNRNYEILDADHLTVASGY-ESLS 180
165 EPLXGEMALPTKLPDLRIAPDIE-VYSKGSRPERDVIYIAVKTDDGDEVFLTAEGK 223
181 LFLGEGVYV--LPLKVLSDIDECVAGNGPSSKCDPIIGI-----GNTLMLGDS 231
224 DDRKPI-----REFVEYKRY-----DPDIIIVGNNHFMWPLYLR 259
222 YASQDIFCLKETTGPVGNVWMEYETKEKELLESKKIPMLDPAVGVNKGPDIFILS 291
260 RARILGK---LDVTRVGAEPFTSVGH-----VSPGRLANDLYDAEMPEI 306
292 RGEILGIESFVSGSKKAKATRTDITWSSNMFSSITTEVEIDGLIIDMMVYIRDPFL 351
307 KIKLSLEVAYLVGNKKSERVIINMWEIPDYDDPKRPLLYOARDV---ATYGLAE 363
352 RSVSLNSVS--IHFLKEQK-----DVPYSISIGELQSKNKTERRIASYCLAD 397
364 KILPFAI-----QLSVYTG.LPDOVGMSVGRFLMXYLIRAFKMKELVPRVYE 412
398 TVLPLRLFNTLVNLYNTELSRVTGPIEYFFTRGAIKIFLTVYRAASKEDFVIPD-ID 456
413 --RPEETVYRATVLEPLAGVHEN--IAVLDPSSMYPMIMIKVNGDITLVREGKCGEGC 469
457 PFESNKTPEGFIIEPRKGFYNKRVSMVDSLSYPSIMLSHNL-----C 500

Qy 470 WEAEVKKRFR-----RCPGFFKTVLER-----LLELRKRVRAEMKKYPPDSP 513
 Db 501 YTTLLTKEQYVILGGTKPTGNYFCSARKKGLPRITLDTLTKRKEELER-EDDSA 559
 Qy 514 EYRLDEROKALKVLANASGYMGSGARWYCRECAKAVTAGRHLLI---RTAI--NIAR 568
 Db 560 LRACINGKQLAFKICANSLYFTGASRGKLPCEFIQSQVTFGFEMLITLTKLIEENSR 619
 Qy 569 KLGLK---VYGDPTDSLFFVYDDPEKVENFIKIKKEELG-----EIKKKYKRL 615
 Db 620 KNGTHDSVYVIGDTSVWVPFDEODIEKVFMSKEISEFTTSFKVPSLEPFKYVYPY 679
 Qy 616 FFEAKKRYAGLL-----EDGRIDIVGFPAVRGDMCELAKEVQKVEIVLKTSEVNKAV 670
 Db 680 LLIN-KKRYALILVSNPNSPKIDTRGIEYTRNDRCRLVKVEVLEMLIYOKNVEKAK 738
 Qy 671 EYRKIVKELEEGKVPLEKVIYIMKTLSTRLEBYTTEADPHVAARKMLSAGYRVSP--GDK 728
 Db 739 QFVDAVADLVIGRTDLSLTVISKSLTKVAGDKESKQAHVOLAEXLRKRDESTAPVLGDR 798
 729 IGVYIV---KGGRIISQAMPYEMWKDPSQIDVTYVVDHQLIPAAARI 773
 Db 799 VPIYIVREKGAHAHEKSEDPVYVLENNLPIDTEYIYSQISKPLSRI 846

RESULT 15

Q9HRV9 PRELIMINARY; PRT; 901 AA.
 AC Q9HRV9;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE DNA polymerase B1.
 GN POLB1 OR VNG0521G.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20504483; PubMed=11016950;
 RA NG W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Laaky S.R., Baliga S.R., Thorsen V., Sirogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weili R., Goo Y.A.,
 RA Liebhauer B., Keller K., Cruz R., Danon M.J., Hough D.W., Dale H.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Jendberger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
 "Genome sequence of Halobacterium species NRC-1".
 Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 RL -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
 CC CC
 CC PYROPHOSPHATE + DNA(N).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC EMBL; AEO05004; AAG19049.1; -
 DR InterPro; IPR002064; DNA_pol_B.
 DR Pfam; PF00136; DNA_pol_B_2_-
 DR Pfam; PF03104; DNA_pol_B-exo; 1.
 DR PRINTS; PR00106; DNAPOLB.
 DR SMART; SM00486; POLB; 1.
 DR PROSITE; PS00116; DNA POLYMERASE B; 1.
 DR DNA replication; DNA-binding; DNA-difected DNA polymerase;
 KW Complete proteome.
 SQ SEQUENCE 901 AA; 101987 MW; 2134704B7AB5E49D CRC64;

Query Match 17.2%; Score 718.5; DB 17; Length 901;
 Best Local Similarity 30.1%; Pred. No. 4.2e-39;
 Matches 239; Conservative 118; Mismatches 293; Indels 145; Gaps 31;

Qy 3 EVFVTVLDDSSIEVVG--KEPQVITIGLNGERVVLLIDRSFRPYVALLAPGADPKOVAQ 60
 Db 58 ELMTQVQVYITTEGSDRBRPVLAVFGRTTDEAEHVRVHGFRPYVAPATLALSD----- 112

Qy 61 RIRALSPKSPITIGVEDDK-----KYFGR--PRVRLIRFVLPFAVREVELVKNVD 111
 Db 113 -----DLTDDVITGTEDGESINGEALTKIFGRTPDVGNIRAF----- 152
 Qy 112 GVEDVLEADIRFAMRYLIDHDLFFPTWYVEAEPLBNKNGFRVDKYYLVKSREPLYGEA 171
 Db 153 ---DHVEADILFNNRLIDKDI--TDGVVPAR-----RADGALVYHHEIAACEV 199
 Qy 172 LAPTKLPDLILAFDIEVYSKQSPRPERRPVLT----- 206
 Db 200 AA-----DLRVNTEFVNDNRGFPEDGEPVVCILTSHDSYRDEVAMLYAAPDATVDA 254
 Qy 207 -AVK-----TDGDD-----EVLFLAEKGDKRPRIEFVYKRVDPDIIIVGNNNHFDMPY 257
 Db 255 TAVPGDPLTDADIDRVVFTTEAHND-----AFISTYEDTNPDVLTGNNPDPFAPYL 309
 Qy 258 LRRARILGIR-----LDVTR--RVGAEPITSHVGHVSPGRANVD--LYDYAE--EMPEIKIK 309
 Db 310 IDRLDELDPRTDHDLDSDRSRV--SEVWTSQWGPVKGKRVFVLDLYAVQRTKYSLEDSY 368
 Qy 310 SLEBVAEY--LGWKKSERVLIINWWEIPDYWDPKKRPLLIQYARDVRYGLAEK--IL 366
 Db 369 RLDAVGEQELGVGK--ERYP---GDIGDLMEDDPER--LLEVMRLVVELCVEIDRKQSIY 421
 Qy 367 PPAIQLSYVTGLPLDQGANVSGFRLEWYLIRAFPKKELVPRVREPRETYGATVLEP 426
 Db 422 AFWDARKLVGCKLED--ATTPGDADVMTVLHAKFG--NFLPSPKQODAEFEGANFDP 478
 Qy 427 LRGVHENIAVLDFSSMYPNIMIKYVGPDTLVPRGCKGCGCEAPEVGRFRRCPPGF 486
 Db 479 ITGVRENVSVLDKSLYPMQMTVINSPETKVDPEFDGD--TYRPTFGNH--FRKSPDGI 535
 Qy 487 FKVYLERLELRKVRPAEMKKYPPDSPETRYLDEROKALKVLANASGYMGSGARWYCR 546
 Db 536 IREWDELLTERREKRYARBDHDEPSEDERYDROQAAVYVIMNSLYGVFGMDRFLYDR 595
 Qy 547 ECAKAVTAMGRHLIRTAIINARLGLKVIYGDPTSLFV-----TYDEKYE 592
 Db 596 AMSAGVSTNREVIDFTEQAABERFGEVAGDPTDSVWLEIGDDMTAALAESFDIEDHI 655
 Qy 593 N--FIKIKKEEL-----GFEIKLEKYKRLFTEAKKRYAG--LLEBGR-----IDIVGFE 639
 Db 656 NAAVDEPARQGLAADHRPQIEFEKLYRRFPQAGKKRVAIGHIVMKEGQVDDVDITGFE 715
 Qy 640 AVRGDMCELAKEVQTKVEIVLKTSEVNKAVERYRKIVKELEEGKVPLEKVIYIMKTLSCR 699
 Db 716 YORSIDIPIYKRVQKEVIDLVREGDVDAVEEYVHGVEFQEGDADLDIDIGIPGIGKR 775
 Qy 700 LEEYTTAAPHVAAK 714
 Db 776 LDNYETDTAQRGAK 790

Search completed: April 23, 2003, 07:31.13
 Job time : 47 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:28:09 ; Search time 42 Seconds

2547.625 Million cell updates/sec

Title: US-10-034-849-2

Sequence: 1 MTEVVFVLDDSYEVGKEP.....LKASATGQKTLFDLAKSK 803

Scoring table: BLOSUM62

arched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_101002:4

1:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT *
2:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT *
3:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT *
4:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT *
5:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT *
6:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT *
7:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT *
8:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT *
9:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT *
10:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT *
11:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT *
12:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT *
13:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT *
14:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT *
15:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT *
16:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT *
17:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT *
18:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT *
19:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT *
20:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT *
21:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT *
22:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT *
23:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4188	100.0	803	20	AAV00936	P. fumarius DNA po
2	4188	100.0	803	23	AAE22118	Pyrobolus fumaria
3	4164	99.4	803	23	AAE22119	Pyrobolus fumaria
4	4122	98.4	799	23	AAE22120	Pyrobolus fumaria
5	3033.5	72.4	803	17	AAE92523	Pyrodicticum occultu
6	3028.5	72.3	803	15	AAE60882	P. occultum DNA pol
7	2963.5	70.8	803	15	AAE60882	P. abyssii DNA poly
8	2129.5	50.8	785	21	AAE19530	Pyrobaculum island
9	1581	37.8	785	21	AAE52022	A. fulgidus AF0497
10	1581	37.8	781	21	AAV51651	A. fulgidus AF0497

11	1563	37.3	788	20	AAE00937
12	1563	37.3	788	21	AAE22114
13	1248.5	29.8	774	23	AAE33416
14	1246.5	29.8	774	23	AAE377018
15	1245.5	29.7	775	23	AAE33417
16	1244.5	29.7	775	23	AAE377019
17	1244.5	29.7	775	23	AAE377013
18	1242.5	29.7	774	23	AAE33402
19	1242.5	29.7	774	23	AAE33415
20	1242.5	29.7	774	23	AAE33415
21	1241.5	29.6	774	23	AAE34514
22	1241.5	29.6	774	23	AAE34514
23	1241.5	29.6	775	19	AAE77017
24	1239.5	29.6	773	23	ABE09843
25	1238.5	29.6	774	19	AAE44731
26	1238.5	29.6	774	19	AAE44732
27	1237.5	29.5	774	19	AAE41313
28	1237.5	29.5	775	17	AAE60255
29	1236.5	29.5	775	17	AAE03687
30	1236.5	29.5	775	17	AAE36733
31	1236.5	29.5	775	20	AAE377282
32	1236.5	29.5	775	20	AAE35677
33	1236.5	29.5	844	23	AAE15887
34	1235.5	29.5	773	19	AAE68818
35	1234.5	29.5	773	19	AAE33154
36	1234.5	29.5	774	18	AAE20048
37	1233.5	29.5	774	18	AAE33111
38	1233.5	29.5	774	19	AAE84455
39	1232.5	29.4	774	19	AAE84457
40	1231.5	29.4	774	19	AAE84450
41	1230.5	29.4	774	19	AAE84454
42	1230.5	29.4	774	19	AAE84455
43	1230.5	29.4	774	19	AAE84455
44	1230.5	29.4	774	19	AAE84457
45	1230.5	29.4	774	19	AAE84455

ALIGNMENTS

[illegible]

XX	RESULT 1
XX	AAV00936
ID	AAV00936 standard; Protein; 803 AA.
XX	
AC	AAV00936;
XX	
DT	02-JUN-1999 (first entry)
XX	
DE	P. fumarius DNA polymerase 1PY2 protein sequence.
XX	
XX	DNA polymerase; thermophilic bacteria; DNA synthesis.
XX	
OS	Pyrolobus fumarius.
XX	
PN	WO907837-A1.
XX	
PD	18-FEB-1999.
XX	
PF	06-AUG-1998; 98WO-US17152.
XX	
PR	06-AUG-1997; 97US-0907166.
XX	
PA	(DIVE-) DIVERSA INC.
XX	
PI	Callen W, Mathur EJ;
XX	
DR	WPI; 1999-180490/15.
XX	
DR	N-PSDB; AAX27283.
XX	
FT	DNA polymerases from extremely thermophilic bacteria - useful for
XX	DNA synthesis
PS	Claim 1; Fig 2; 72PP; English.

XX This sequence is a DNA polymerase of the invention, that was
 CC isolated from a thermophilic bacteria. The polymerases are used in DNA
 CC synthesis and as immunogens to raise antibodies (useful for affinity
 CC purification and to screen for related enzymes). Fragments of the DNA
 CC full-length polymerases are used as probes to isolate related or
 CC polymerases catalyze DNA synthesis by the addition of deoxynucleotides to
 CC the 3' end of a polynucleotide chain, using a complementary
 CC polynucleotide strand as a template. The polymerases have optimum
 CC activity at over 60 deg. C and can renature and regain activity after
 CC exposure to temperatures above 70 deg. C.

XX Sequence 803 AA;

Query Match 100.0%; Score 4188; DB 20; Length 803;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEVFTVLDSSYEYVGEKPEQVIYIWIAGNGBRVLLIDRSFRPYFALLAPGADPKQVQAQ 60
 1 MTEVFTVLDSSYEYVGEKPEQVIYIWIAGNGBRVLLIDRSFRPYFALLAPGADPKQVQAQ 60
 QY 61 RIRALSRRKSPPIIGVEDDKRKYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 61 RIRALSRRKSPPIIGVEDDKRKYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 Db 61 RIRALSRRKSPPIIGVEDDKRKYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 QY 121 IRFAMRYLIDHDLPEFTMYRVEAEPLLENKMGFRVDKXYLYVKSPPELVGEALAPTKLPDL 180
 121 IRFAMRYLIDHDLPEFTMYRVEAEPLLENKMGFRVDKXYLYVKSPPELVGEALAPTKLPDL 180
 Db 121 IRFAMRYLIDHDLPEFTMYRVEAEPLLENKMGFRVDKXYLYVKSPPELVGEALAPTKLPDL 180
 QY 181 RILAFDIEVYSKQSSPRPRDPVYIYAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 181 RILAFDIEVYSKQSSPRPRDPVYIYAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 Db 181 RILAFDIEVYSKQSSPRPRDPVYIYAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 QY 241 PDIIIVGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVLDYDA 300
 241 PDIIIVGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVLDYDA 300
 Db 241 PDIIIVGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVLDYDA 300
 QY 301 EEMPEIKIKSLSEVAEYLGVMKKSSRVLIINWMEIPDWDPPKRPPLLQYAADVRATYTG 360
 301 EEMPEIKIKSLSEVAEYLGVMKKSSRVLIINWMEIPDWDPPKRPPLLQYAADVRATYTG 360
 Db 301 EEMPEIKIKSLSEVAEYLGVMKKSSRVLIINWMEIPDWDPPKRPPLLQYAADVRATYTG 360
 QY 361 LAEKILPFAIQLSVTGLPLDOYGAMSVGFLEWYLIRAAFKKKEIVPRKVRPRETYRG 420
 361 LAEKILPFAIQLSVTGLPLDOYGAMSVGFLEWYLIRAAFKKKEIVPRKVRPRETYRG 420
 Db 361 LAEKILPFAIQLSVTGLPLDOYGAMSVGFLEWYLIRAAFKKKEIVPRKVRPRETYRG 420
 QY 421 AIVLEPLRGVHENIYAVLDFSSMYPNIMIKYVNGPDTLVRRGEKCGCGCEGMEAPVKAHFR 480
 421 AIVLEPLRGVHENIYAVLDFSSMYPNIMIKYVNGPDTLVRRGEKCGCGCEGMEAPVKAHFR 480
 Db 421 AIVLEPLRGVHENIYAVLDFSSMYPNIMIKYVNGPDTLVRRGEKCGCGCEGMEAPVKAHFR 480
 QY 481 RCPGPFKTYLERLLERKRVRAEMKCTPPSPPEYRLIDDEROKALVLANASYGMWSG 540
 481 RCPGPFKTYLERLLERKRVRAEMKCTPPSPPEYRLIDDEROKALVLANASYGMWSG 540
 Db 481 RCPGPFKTYLERLLERKRVRAEMKCTPPSPPEYRLIDDEROKALVLANASYGMWSG 540
 QY 541 ARMYGRCACAVTAMGRHLIRTAIINARKLGKVIYGDTSLFVYDPEKVENFKIIKE 600
 541 ARMYGRCACAVTAMGRHLIRTAIINARKLGKVIYGDTSLFVYDPEKVENFKIIKE 600
 Db 541 ARMYGRCACAVTAMGRHLIRTAIINARKLGKVIYGDTSLFVYDPEKVENFKIIKE 600
 QY 601 ELGFEIKLEKYKRLPFTTEAKRGVAGLLEDSRIDVFEAVRGDMCELAKEVQTQVEIV 660
 601 ELGFEIKLEKYKRLPFTTEAKRGVAGLLEDSRIDVFEAVRGDMCELAKEVQTQVEIV 660
 Db 601 ELGFEIKLEKYKRLPFTTEAKRGVAGLLEDSRIDVFEAVRGDMCELAKEVQTQVEIV 660
 QY 720 LKTSVNVKAVEYVRKIYVKELEGKVPIDKLVMTKLSRLSEETTEPYNVAAKMLTAG 720
 720 LKTSVNVKAVEYVRKIYVKELEGKVPIDKLVMTKLSRLSEETTEPYNVAAKMLTAG 720
 Db 720 LKTSVNVKAVEYVRKIYVKELEGKVPIDKLVMTKLSRLSEETTEPYNVAAKMLTAG 720
 QY 780 YRSPGPKIGVYVKGSGRIISORAMPYMWDPDSQIVTYVVOHIIIPALNIIIGYRIT 780
 780 YRSPGPKIGVYVKGSGRIISORAMPYMWDPDSQIVTYVVOHIIIPALNIIIGYRIT 780
 Db 780 YRSPGPKIGVYVKGSGRIISORAMPYMWDPDSQIVTYVVOHIIIPALNIIIGYRIT 780
 QY 803 EKKLKASATGOKTLFPLAKSK 803
 803 EKKLKASATGOKTLFPLAKSK 803
 Db 803 EKKLKASATGOKTLFPLAKSK 803

Db 781 EKKLKASATGOKTLFPLAKSK 803

RESULT 2

ID. AAE22118 standard; Protein; 803 AA.

AAE22118;

25-JUL-2002 (first entry)

Pyrolobus fumaria DNA polymerase, 1PY2.

DNA polymerase; thermostable; enzyme.

Pyrolobus fumaria.

MO200220735-A2.

14-MAR-2002.

06-SEP-2001; 2001MO-US28007.

06-SEP-2000; 2000US-0656309.

(DIVE-) DIVERSA CORP.

Callen W, Mathur EJ, Short JM,

WPI; 2002-362247/39.

N-PDB; AAD35187.

New thermostable polymerase useful for sequencing DNA, amplifying

double stranded DNA, or incorporating a non-natural nucleotide or a

nucleotide analog into a DNA molecule

Claim 42; Fig 1A-E; 16tp; English.

The invention relates to thermostable DNA polymerases having high

temperature polymerase activity, such as those derived from

Pyrolobus fumaria and nucleic acid molecules encoding such polymerases.

Polymerases are useful for catalyzing the formation or repair of a

nucleic acid sequence and for modifying small molecules. They are also

useful for sequencing DNA molecules, for preparing cDNA from mRNA, for

amplifying double stranded DNA molecules and for incorporating non-

natural nucleotides or nucleotide analogues into a DNA molecule. The

present sequence is Pyrolobus fumaria DNA polymerase.

Sequence 803 AA;

Query Match 100.0%; Score 4188; DB 23; Length 803;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEVFTVLDSSYEYVGEKPEQVIYIWIAGNGBRVLLIDRSFRPYFALLAPGADPKQVQAQ 60
 1 MTEVFTVLDSSYEYVGEKPEQVIYIWIAGNGBRVLLIDRSFRPYFALLAPGADPKQVQAQ 60
 Db 1 MTEVFTVLDSSYEYVGEKPEQVIYIWIAGNGBRVLLIDRSFRPYFALLAPGADPKQVQAQ 60
 QY 61 RIRALSRRKSPPIIGVEDDKRKYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 61 RIRALSRRKSPPIIGVEDDKRKYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 Db 61 RIRALSRRKSPPIIGVEDDKRKYFGRRPRVLRIRTVLPEAVREYRELKVNVDGVEVLEAD 120
 QY 121 IRFAMRYLIDHDLPEFTMYRVEAEPLLENKMGFRVDKXYLYVKSPPELVGEALAPTKLPDL 180
 121 IRFAMRYLIDHDLPEFTMYRVEAEPLLENKMGFRVDKXYLYVKSPPELVGEALAPTKLPDL 180
 Db 121 IRFAMRYLIDHDLPEFTMYRVEAEPLLENKMGFRVDKXYLYVKSPPELVGEALAPTKLPDL 180
 QY 181 RILAFDIEVYSKQSSPRPRDPVYIYAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 181 RILAFDIEVYSKQSSPRPRDPVYIYAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 Db 181 RILAFDIEVYSKQSSPRPRDPVYIYAVKTDGDEVLFIAEGKODRKPIREFVEYVKRYD 240
 QY 241 PDIIIVGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVLDYDA 300
 241 PDIIIVGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVLDYDA 300
 Db 241 PDIIIVGYNNHFDMPYLLRRARILGKLDVTRVGAEPPTSVAHGVSVEGRNLVLDYDA 300

QY 301 BEMPEIKIKSLAEVAEYLVGKKSSERVITNMWEIPDYWDDPKRPLLIQVARDVDRATYG 360
 DB 301 BEMPEIKIKSLAEVAEYLVGKKSSERVITNMWEIPDYWDDPKRPLLIQVARDVDRATYG 360
 QY 361 LAEKILPFAIQLSVYVGLPLDQVGMASVGRLEWYLIRAAFKMKELVNNRVERPEETRYG 420
 DB 361 LAEKILPFAIQLSVYVGLPLDQVGMASVGRLEWYLIRAAFKMKELVNNRVERPEETRYG 420
 QY 421 AIVLEPLRGVHENIAVLDFSSMYPNIMIKYVGPDTLVRPGEKCGECGCEAPVKKRFR 480
 DB 421 AIVLEPLRGVHENIAVLDFSSMYPNIMIKYVGPDTLVRPGEKCGECGCEAPVKKRFR 480
 QY 481 RCPGPFKTYLERLLELRKVRAEKKYPPDSPERYLLDEROKALKYLANASYGMGWSG 540
 DB 481 RCPGPFKTYLERLLELRKVRAEKKYPPDSPERYLLDEROKALKYLANASYGMGWSG 540
 QY 541 ARMYCRECAKAVTAMGRHLIRTAINTARKLGLKVIYGDTSLPVYDPEKVENFIKIIE 600
 DB 541 ARMYCRECAKAVTAMGRHLIRTAINTARKLGLKVIYGDTSLPVYDPEKVENFIKIIE 600
 QY 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTKVEIV 660
 DB 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTKVEIV 660
 QY 661 LKTSVNKAVEYVRKIVKELEEGKVPLEKIYIMKTLSKRLSEYTTAEAPHVAAKRMISAG 720
 DB 661 LKTSVNKAVEYVRKIVKELEEGKVPLEKIYIMKTLSKRLSEYTTAEAPHVAAKRMISAG 720
 QY 721 YRVSFGDKIGVIVYKGGRIISORAMPYFMVKDPSQIDVTYVVDHQLIPALRLIGYFGIT 780
 DB 721 YRVSFGDKIGVIVYKGGRIISORAMPYFMVKDPSQIDVTYVVDHQLIPALRLIGYFGIT 780
 QY 781 EKKLKASATGQKTLFDPLAKSK 803
 DB 781 EKKLKASATGQKTLFDPLAKSK 803
 RESULT 3
 AAE22119
 ID AAE22119 standard; Protein; 803 AA.
 XX AC AAE22119;
 XX DT 25-JUL-2002 (first entry)
 XX Pyrolobus fumaria DNA polymerase, PLF1831.
 DNA polymerase; thermostable; enzyme.
 XX OS Pyrolobus fumaria.
 XX PN W0200220735-A2.
 XX PD 14-MAR-2002.
 XX PF 06-SEP-2001; 2001WO-US28007.
 XX PR 06-SEP-2000; 2000US-0656309.
 XX PA (DIVE-) DIVERSA CORP.
 XX PI Calien W, Mathur EJ, Short JM;
 XX DR WPI; 2002-362247/39.
 XX PT N-PSDB; AAD35188.
 XX New thermostable polymerase useful for sequencing DNA, amplifying
 PT double stranded DNA, or incorporating a non-natural nucleotide or a
 PT nucleotide analog into a DNA molecule
 XX PS Claim 42; Fig 11; 161p; English.
 XX

CC The invention relates to thermostable DNA polymerases having high
 CC temperature polymerase activity, such as those derived from
 CC Pyrolobus fumaria and nucleic acid molecules encoding such polymerases.
 CC Polymerases are useful for catalysing the formation or repair of a
 CC nucleic acid sequence and for modifying small molecules. They are also
 CC useful for sequencing DNA molecules, for preparing cDNA from mRNA, for
 CC amplifying double stranded DNA molecules and for incorporating non-
 CC natural nucleotides or nucleotide analogues into a DNA molecule. The
 CC present sequence is Pyrolobus fumaria DNA polymerase.

XX Sequence 803 AA;

Query Match 99.4%; Score 4164; DB 23; Length 803;

Best Local Similarity 99.3%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 797; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTEVFTVLDSSEYVGEKEPQVITWGIAENGERVILDRSFRPFYFALLAPGADPKQVAQ 60
 DB 1 MTEVFTVLDSSEYVGEKEPQVITWGIAENGERVILDRSFRPFYFALLAPGADPKQVAQ 60
 QY 61 RIRALSRPKSPIIGVEDDKKRYGRRPVLRIRTVLPEAVREYRELVKNVGVEDVLEAD 120
 DB 61 RIRALSRPKSPIIGVEDDKKRYGRRPVLRIRTVLPEAVREYRELVKNVGVEDVLEAD 120
 QY 121 IRFAMRYLIDHDLPPFTWYVVEABPLENKMGPVVDKVIYKSRPEPLYGALAPTKLPDL 180
 DB 121 IRFAMRYLIDHDLPPFTWYVVEABPLENKMGPVVDKVIYKSRPEPLYGALAPTKLPDL 180
 QY 181 RILAFDIEVYSKQSPRRERDPVIVIAVKTDDGDEVLFIAEGDODRKPILFEFVYVRYD 240
 DB 181 RILAFDIEVYSKQSPRRERDPVIVIAVKTDDGDEVLFIAEGDODRKPILFEFVYVRYD 240
 QY 241 PDIIYVGNHHPWYLLRPARILGILKLDVTRRGVAPPTTSVHGHSVPLRLNVDLYDVA 300
 DB 241 PDIIYVGNHHPWYLLRPARILGILKLDVTRRGVAPPTTSVHGHSVPLRLNVDLYDVA 300
 QY 301 BEMPEIKIKSLAEVAEYLVGKKSSERVITNMWEIPDYWDDPKRPLLIQVARDVDRATYG 360
 DB 301 BEMPEIKIKSLAEVAEYLVGKKSSERVITNMWEIPDYWDDPKRPLLIQVARDVDRATYG 360
 QY 361 LAEKILPFAIQLSVYVGLPLDQVGMASVGRLEWYLIRAAFKMKELVNNRVERPEETRYG 420
 DB 361 LAEKILPFAIQLSVYVGLPLDQVGMASVGRLEWYLIRAAFKMKELVNNRVERPEETRYG 420
 QY 421 AIVLEPLRGVHENIAVLDFSSMYPNIMIKYVGPDTLVRPGEKCGECGCEAPVKKRFR 480
 DB 421 AIVLEPLRGVHENIAVLDFSSMYPNIMIKYVGPDTLVRPGEKCGECGCEAPVKKRFR 480
 QY 481 RCPGPFKTYLERLLELRKVRAEKKYPPDSPERYLLDEROKALKYLANASYGMGWSG 540
 DB 481 RCPGPFKTYLERLLELRKVRAEKKYPPDSPERYLLDEROKALKYLANASYGMGWSG 540
 QY 541 ARMYCRECAKAVTAMGRHLIRTAINTARKLGLKVIYGDTSLPVYDPEKVENFIKIIE 600
 DB 541 ARMYCRECAKAVTAMGRHLIRTAINTARKLGLKVIYGDTSLPVYDPEKVENFIKIIE 600
 QY 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTKVEIV 660
 DB 601 ELGFEIKLEKYKRLFFTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVQTKVEIV 660
 QY 661 LKTSVNKAVEYVRKIVKELEEGKVPLEKIYIMKTLSKRLSEYTTAEAPHVAAKRMISAG 720
 DB 661 LKTSVNKAVEYVRKIVKELEEGKVPLEKIYIMKTLSKRLSEYTTAEAPHVAAKRMISAG 720
 QY 721 YRVSFGDKIGVIVYKGGRIISORAMPYFMVKDPSQIDVTYVVDHQLIPALRLIGYFGIT 780
 DB 721 YRVSFGDKIGVIVYKGGRIISORAMPYFMVKDPSQIDVTYVVDHQLIPALRLIGYFGIT 780
 QY 781 EKKLKASATGQKTLFDPLAKSK 803
 DB 781 EKKLKASATGQKTLFDPLAKSK 803

Best Local Similarity 70.8%; Pred. No. 1.6e-276;
Matches 571; Conservative 106; Mismatches 121; Indels 7; Gaps 5;

```

Qy 1 MTEVY-FTVLSSYEVNGKEPQVIMWIAENGSRVYLIDRSFRPFYALLAPGADP--KQ 57
Db 1 MTEIIEFVLDDSYEILGKEPVILMGITLDGRVLLDHRFRPYFALLARGEDWEE 60
Qy 58 VAORIRALSRPKSPIIGVEDDKRYFGRRRLIRITVLPFAVRREYRLVKNVGVEDVL 117
Db 61 IASIRRLSVKSPILIDAKPLDKRFGPRKAVKITMIPESVRYRAVAKKIEGEDSL 120
Qy 118 EADIRFAMRYLIDHDLFPFTYRVVEAPLENKMGFRVDKYLKVSRRPPLYGEALAPTKL 177
Db 121 EADIRFAMRYLIDKRLYFPTVYRIPEVDAGRNPGRVDRVYKVGADDEPL--ADITRIDL 178
Qy 178 PDILIAFDIEVYSGKQSPREPDVYIAVKTDDGDEVLPFAEKDQRKPIREVEYVK 237
Db 179 PPMRLVAFDIEVYSGRSPNARDPVIIIVSLRDESEKRLIEAGHDHRRVLRFEVEYR 238
Qy 238 RYDPDIIYGNNNHFDMPYLLRRARILGKIDVTRRGAEPFTSVHGHVSGRLVNDLY 297
Db 239 AFDDIIVGYNHFDMPYLLMERARRLGKIDVTRRGAEPFTSVYGHVSVQGRVNDLY 298
Qy 298 DYAEEMPEIKIKLSEVAEYLVGMKKSERYIINWMEIPDYDDPKRPLDQYARDVRA 357
Db 299 DYAEEMPEIKIKLSEVAEYLVGMKKSERYIINWMEIPDYDDPKRPLDQYARDVRA 358
Qy 358 TYGLAEKILPFAIQLSTVTGLPLDQVGMVSGFRLEWYLIRAAFKMELVNNRVERPET 417
Db 359 TYGLAEKILPFAIQLSTVTGLPLDQVGMVSGFRLEWYLIRAAFKMELVNNRVERPET 418
Qy 418 YRGAIVLEPLRGVHENIAVLDPSSMYPNIMIKYVNGPDTLVRPGEKGC--GCWEAPEVK 476
Db 419 YKGAIVLKLKGVHENVAVLDPSSMYPNIMIKYVNGPDTLVDPSCEKYGCTVAPEVG 478
Qy 477 HRFRCPGPFKTVLERLLERKRVRAEMKKYPPDSEPRYLIDRQKALKVLANASYGM 536
Db 479 HRFRSPGPFKTVLERLLERKRVRAEMKKYPPDSEPRYLIDRQKALKVLANASYGM 538
Qy 537 GWSGARWYCRCAKAVTAMGRHLIRTAIRAKGLKTYIGDTDSLPTTYPEKYNFIK 596
Db 539 GWSGARWYCRCAKAVTAMGRHLIRTAIRAKGLKTYIGDTDSLPTTYPEKYNFIK 598
Qy 597 IIEKELGEIELEKYKRLFTFEAKRYAGLLEDRIDIVGEAVRGDMCELAKEVOTKV 656
Db 599 FVEKELGEIEIDIKYKRVFTFEAKRYAGLLEDRIDIVGEAVRGDMCELAKEVOTKV 658
Qy 657 VEIYLTSEVNAKAVEYKIVKELLEGKVPTEKLVIMKTLSEKLEEYTTTEAPHVAAKRM 716
Db 659 AEIYLTNGVNDKALSYIREVILKOLREGKVPITKLIIMKTLSEKLEEYTHDAPHVMAARM 718
Qy 717 LSAARVSPGKIGVIVYKGGSRISQRAPIPMYKQDSQIDVTYVHDHIIIPALRIIGY 776
Db 719 KEAEVYSPGKIGVIVYKGGSRISQRAPIPMYKQDSQIDVTYVHDHIIIPALRIISY 777
Qy 777 FGITEKKLKASATGOKTLFDPLAKK 801
Db 778 FGITEKKLKAAATVORSLEDFPASK 802

```

RESULT 6
AAR60883
ID AAR60883 standard; Protein: 803 AA.

AC AAR60883;
XX 19-JUL-1995 (first entry)
DT
XX
DE P. occultum DNA polymerase.
KM Polymerase chain reaction; PCR; amplify; primer; probe;
KW Pyrodictum occultum; DNA polymerase; P. abyssi; condensation;
DNA sequencing; nick translation; reverse transcription;

KM heat stability; heat resistant polymerase; DNA helix stabiliser.

XX Pyrodictum occultum.

XX EP624641-A.

XX 17-NOV-1994.

XX 02-MAY-1994; 94EP-0106811.

XX 14-MAY-1993; 93US-0062368.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Gelfand DH, Wang AM;

XX WPI; 1994-350781/44.

XX N-PDB; AAQ73844.

PT New heat stable nucleic acid polymerase from Pyrodictum species

PT - and relates DNA, vectors and transformed hosts, resistant to

PT denaturation at 100 deg C, useful for nucleic acid amplification,

PT in DNA sequencing etc.

PS Claim 10; Page 33-35; 45pp; English.

CC This sequence represents the Pyrodictum occultum DNA polymerase gene.

CC The P. occultum DNA polymerase gene shows regions of homology to the

CC P. abyssi DNA polymerase genes, and the primers given in AAQ73824-25

CC are based on these regions. The amplified DNA encodes this DNA

CC polymerase which catalyses condensation of dNTPs to form complementary

CC strand DNA. This polymerase may be used in nucleic acid amplification,

CC partic. PCR, but also for DNA sequencing, nick translation and reverse

CC transcription. The polymerase has better heat stability than known

CC heat resistant polymerases. Temperatures of 100 deg. C can be

CC tolerated, facilitating strand separation and elimination of the

CC need for DNA helix stabilisers.

CC Sequence 803 AA;

Qy Query Match 72.3%; Score 3028.5; DB 15; Length 803;

Qy Best Local Similarity 70.8%; Pred. No. 4.6e-276;

Qy Matches 570; Conservative 106; Mismatches 122; Indels 7; Gaps 5;

Qy 1 MTEVY-FTVLSSYEVNGKEPQVIMWIAENGSRVYLIDRSFRPFYALLAPGADP--KQ 57

Qy 1 MTEIIEFVLDDSYEILGKEPVILMGITLDGRVLLDHRFRPYFALLARGEDWEE 60

Qy 58 VAORIRALSRPKSPIIGVEDDKRYFGRRRLIRITVLPFAVRREYRLVKNVGVEDVL 117

Qy 61 IASIRRLSVKSPILIDAKPLDKRFGPRKAVKITMIPESVRYRAVAKKIEGEDSL 120

Qy 118 EADIRFAMRYLIDHDLFPFTYRVVEAPLENKMGFRVDKYLKVSRRPPLYGEALAPTKL 177

Qy 121 EADIRFAMRYLIDKRLYFPTVYRIPEVDAGRNPGRVDRVYKVGADDEPL--ADITRIDL 178

Qy 178 PDILIAFDIEVYSGKQSPREPDVYIAVKTDDGDEVLPFAEKDQRKPIREVEYVK 237

Qy 179 PPMRLVAFDIEVYSGRSPNARDPVIIIVSLRDESEKRLIEAGHDHRRVLRFEVEYR 238

Qy 238 RYDPDIIYGNNNHFDMPYLLRRARILGKIDVTRRGAEPFTSVHGHVSGRLVNDLY 297

Qy 239 AFDDIIVGYNHFDMPYLLMERARRLGKIDVTRRGAEPFTSVYGHVSVQGRVNDLY 298

us 5191686R

```

Db 419 YKGAIVLKLPGKHENVVLDSSMYPSIMIKVNGDITVDPPSCPCPKGCGCYAAPEVG 478
Cc 477 HRPFRCPGPFKTYLERLELAKRYRAMKCYPPDSPEYLLDERGKALKYLANASYGM 536
Cc 479 HRPFRSPGPFKTYLERLELAKRYRAMKCYPPDSPEYLLDERGKALKYLANASYGM 538
Cc 537 GMSGARWYCRECAKANTANGRHILITAINIRKGLKVIYIGDTSLEFVYDPEKVENFIK 596
Cc 539 GMSHARWYCRCAEAVTANGRHILITALEYARKGLKVIYIGDTSLEFVYDPEKVENFIK 598
Cc 597 IIEELGFEIKLEKRYKRLFTFEAKKRYAGLLEDGRIDIVGFEAVRGDWCELAKEVQTKV 656
Cc 599 FVEKELEFEIKIDKIKYKVFTEAKKRYVGLLEDGRIDIVGFEAVRGDWCELAKEVQTKV 658
Cc 657 VEIVLKTSEVNKAVEYRKIVLEBEGKVPTEKLVIMKTLSEKLEBYTTEAPHVAAKRM 716
Cc 659 AEIVLNTGAVNDKASIVIREVILQLEBGRVPTKLIIMKTLSEKLEBYTTEAPHVAAKRM 718
Cc 717 LSAGRYVSPGDKIGYIVVKGGRISORAMPYFMVADPQIDVTYVYDHOIIPALRILGY 776
Cc 719 KENGAVYSPGDKIGYIVVKGSGVSRAVYFMV-DSTIDVYIYDHOIVPAAALRILSY 777
Cc 777 FGITEKKLKASATGOKTLFPFLAKK 801
Cc 778 FGYTEKQLKAASATGORSLEDFPFAK 802

RESULT 7
ID AAR60882 standard; Protein; 803 AA.
Cc AAR60882;
Cc 19-JUL-1995 (first entry)
Cc
Cc P. abyssi DNA polymerase.
Cc
Cc Polymerase chain reaction; PCR; amplify; primer; probe;
Cc Pyrodicticum occultum; DNA polymerase; P. abyssi; condensation;
Cc RNA sequencing; nick translation; reverse transcription;
Cc heat stability; heat resistant polymerase; DNA helix stabiliser.
Cc
Cc Pyrodicticum abyssi.
Cc
Cc EPE24641-A.
Cc 17-NOV-1994.
Cc
Cc 02-MAY-1994; 94EP-0106811.
Cc
Cc 14-MAY-1993; 93US-0062368.
Cc
Cc (HOFF) HOFFMANN LA ROCHE & CO AG F.
Cc
Cc Gelfand DH, Wang AM;
Cc
Cc WPI; 1994-350781/44.
Cc
Cc N-PSDB; AAO73843.
Cc
Cc New heat stable nucleic acid polymerase from Pyrodicticum species
Cc - and relates DNA, vectors and transformed hosts, resistant to
Cc denaturation at 100 deg.C, useful for nucleic acid amplification,
Cc in DNA sequencing etc.
Cc
Cc Claim 8; Page 28-31; 45pp; English.
Cc
Cc This sequence represents the Pyrodicticum abyssi DNA polymerase gene.
Cc The P. abyssi DNA polymerase gene shows regions of homology to the
Cc P. occultum DNA polymerase gene, and the primers given in AAO73824-25
Cc are based on these regions. The amplified DNA encodes this DNA
Cc polymerase which catalyses condensation of dNTPs to form complementary
Cc strand DNA. This polymerase may be used in nucleic acid amplification,
Cc

```

```

Cc partic. PCR, but also for DNA sequencing, nick translation and reverse
Cc transcription. The polymerase has better heat stability than known
Cc heat resistant polymerases. Temperatures of 100 deg. C can be
Cc tolerated, facilitating strand separation and elimination of the
Cc need for DNA helix stabilisers.
Cc
Cc Sequence 803 AA;
Cc
Cc Query Match 70.8%; Score 2963.5; DB 15; Length 803;
Cc Best Local Similarity 68.4%; Pred. No. 6.2e-270;
Cc Matches 551; Conservative 118; Mismatches 129; Indels 7; Gaps 5;
Cc
Cc 1 MNEVV-FTVLDSYEVNGKEPQVILINGIENGERVLLIDRSFRPYFALLAPGADPK-Q 57
Cc 1 MEALIEFVLDSYEVNGKEPQVILINGIENGERVLLIDRSFRPYFALLAPGADPK-Q 57
Cc
Cc 58 VQORIRALSRRKSPILIGVEDDKRYGPRPRVLRIRTVLPAVAREVELYKQVDEVDVL 117
Cc 61 VVAAIRRLSMAKSPILIEAKVSKKYPGRPRKAVKVTIVIPESVREYREAVKLEGEVDEL 120
Cc
Cc 118 EADIRPAMRYLIDDLPEFTWVVEAELENMGGRVQXVVLKSPREPIYGEALAPTL 177
Cc 121 EADIRPAMRYLIDDLPEFTWVVEAELENMGGRVQXVVLKSPREPIYGEALAPTL 178
Cc
Cc 178 PDRLIAPDIVYKQSGSPREPDPYIVAVTGDGDEVLEIAGKQDKPIEFEPYK 237
Cc 179 PEKRVLPDIVYKQSGSPREPDPYIVAVTGDGDEVLEIAGKQDKPIEFEPYK 238
Cc
Cc 238 RYDPIIIVGNNHPMPYLLRPARILGKIDVTRVCAEPTTSVHGVSVGRNLNDV 297
Cc 239 SPDPDIVGNSNPNMPYLLRPARILGKIDVTRVCAEPTTSVHGVSVGRNLNDV 298
Cc
Cc 298 DYAEEMPEIKIKSLAEVAVLYGVKKSSEVYIINMEIPDYVDKPKRPLLOVARDVNA 357
Cc 299 NYVEEMHEIKYKTLAEVAVLYGVKKSSEVYIINMEIPDYVDKPKRPLLOVARDVNA 358
Cc
Cc 358 TYGLAEKILPFAIOLSVYVGLPLDQVAMSVGRLEWYLIIRAFKMKELVPRNVERPEET 417
Cc 359 TYGLAEKILPFAIOLSVYVGLPLDQVAMSVGRLEWYLIIRAFKMKELVPRNVERPEET 418
Cc
Cc 418 YKGAIVLKLPGKHENVVLDSSMYPSIMIKVNGDITVDPPSCPCPKGCGCYAAPEVG 476
Cc 419 YKGAIVLKLPGKHENVVLDSSMYPSIMIKVNGDITVDPPSCPCPKGCGCYAAPEVG 478
Cc
Cc 477 HRPFRCPGPFKTYLERLELAKRYRAMKCYPPDSPEYLLDERGKALKYLANASYGM 536
Cc 479 HRPFRSPGPFKTYLERLELAKRYRAMKCYPPDSPEYLLDERGKALKYLANASYGM 538
Cc
Cc 537 GMSGARWYCRECAKANTANGRHILITAINIRKGLKVIYIGDTSLEFVYDPEKVENFIK 596
Cc 539 GMSHARWYCRCAEAVTANGRHILITALEYARKGLKVIYIGDTSLEFVYDPEKVENFIK 598
Cc
Cc 597 IIEELGFEIKLEKRYKRLFTFEAKKRYAGLLEDGRIDIVGFEAVRGDWCELAKEVQTKV 656
Cc 599 FVEKELEFEIKIDKIKYKVFTEAKKRYVGLLEDGRIDIVGFEAVRGDWCELAKEVQTKV 658
Cc
Cc 657 VEIVLKTSEVNKAVEYRKIVLEBEGKVPTEKLVIMKTLSEKLEBYTTEAPHVAAKRM 716
Cc 659 AEIVLNTGAVNDKASIVIREVILQLEBGRVPTKLIIMKTLSEKLEBYTTEAPHVAAKRM 718
Cc
Cc 717 LSAGRYVSPGDKIGYIVVKGGRISORAMPYFMVADPQIDVTYVYDHOIIPALRILGY 776
Cc 719 KENGAVYSPGDKIGYIVVKGSGVSRAVYFMV-DSTIDVYIYDHOIVPAAALRILSY 777
Cc
Cc 777 FGITEKKLKASATGOKTLFPFLAKK 801
Cc 778 FGYTEKQLKAASATGORSLEDFPFAK 802

RESULT 8
ID AAB19530 standard; Protein; 785 AA.
Cc AAB19530
Cc

```

AC AAB19530;
 XX
 DT 09-JAN-2001 (first entry)
 XX
 DE Pyrobaculum islandicum DNA polymerase.
 XX
 KM DNA polymerase; 3'-5' exonuclease; thermostable; DNA amplification;
 KM DNA sequencing.
 XX
 OS Pyrobaculum islandicum.
 XX
 PN WO200053772-A1.
 XX
 PD 14-SEP-2000.
 XX
 PF 04-MAR-2000; 2000WO-EP01917.
 XX
 PR 06-MAR-1999; 99EP-0104539.
 PR 02-SEP-1999; 99EP-0117245.
 (HOF) ROCHE DIAGNOSTICS GMBH.
 XX
 PI Kehler M, Frey B, Sobek H, Antikhan G;
 XX
 DR MPI; 2000-587439/55.
 DR N-PSDB; AAA50678.
 XX
 PT Novel DNA polymerase of Pyrobaculum islandicum used for determining DNA
 PT sequences, comprising six conserved motifs indicative of family B DNA
 PS polymerases and possessing exonuclease activity
 PS
 XX Disclosure; Fig 1/2; 38pp; English.
 CC
 CC The present sequence is that of the novel DNA polymerase of the
 CC anaerobic hyperthermophilic Archaeon, Pyrobaculum islandicum DSM24184.
 CC The sequence was deduced from a DNA polymerase gene (see AAA50678).
 CC isolated from P. islandicum chromosomal DNA. The enzyme includes 6
 CC conserved motifs, indicative of a family B DNA polymerase. A method
 CC for producing recombinant DNA polymerase using Escherichia coli BL21
 CC cells transformed by a vector including the P. islandicum gene is
 CC claimed. The P. islandicum DNA polymerase is thermostable, having a
 CC half-life of 30-40 min at 100 degree C and over 10 hr at 90 degree C.
 CC It can be used for in vitro amplification of DNA fragments and for
 CC DNA sequencing (claimed). The enzyme also has 3'-5' proofreading
 CC exonuclease activity, providing a substantially lower base
 CC incorporation error rate than non-proofreading DNA polymerases.
 CC
 CC Sequence 785 AA;
 Query Match 50.8%; Score 2129.5; DB 21; Length 785;
 Best Local Similarity 53.4%; Pred. No. 2.7e-191;
 Matches 429; Conservative 141; Mismatches 195; Indels 39; Gaps 13;
 QY 9 LDSSYEVNGKEPOVILWIGIAENGERVVLIDRSFRPYF---ATLAPGADPKVQAQRAL 65
 DB 8 LDITYAVVSGVPEIRIRIFILSSGERVVLIDRSFKYFVDCVCEPAA-----LKTAL 60
 QY 66 SRPSPITIGVEDDKKYGFRPRVLRITVLPEAVREYRELWKNVGVEDVLEADIRPAM 125
 DB 61 SR-VAPIDVDQIVERRRFGRSKKFLKVIKIPEDVRKLRKRAMSIPRVSGVYEAIDIRYM 119
 QY 126 RYLLDHDLPFTWYRVVEPELENKMGFRVDKY--YLVSSREPLYG--EALAPTLPLR 181
 DB 120 RYMLDMGVVPCSMVAVEE-----GRLGGIPYVVSQ---WYIDEGFPFS---Lk 166
 QY 182 ILAPDIEVYSKQSGPRPRDPVIVAVKTDGDEVLFAEGKDRKPIREFVEYVKRYP 241
 DB 167 VMAPDIEVYNRGSDPIRDVVMALIKTNGHEVFEASKDRGVRAVDPIRSIDP 226
 QY 242 DIYGVNNHNDWPIYLRRARILGKLDVTRVGAEPFTTSHGHVSVGRINVDLYDAE 301
 DB 227 DVIYGVNSNGPDWPIYLRARAVGVPLKVD-RLSNPQOSVYGHWSIVGRANVDLYNIVE 285

QY 302 EMPREIKISLEEVAEVLYGVNKKSERVILNMWMEIPDYWDDPKKRPPLLQVARDVRAITYGL 361
 DB 286 EPEIKIKTLIDRAVEAIEYGVNKKSERVILPGHKIYETWKKDRKPLRLRYLDDVRSTLGL 345
 QY 362 AEKILPEAIOLSYVTGLPLDOVGAMSGVFRLEWVILRAAFKMEELVNRVEREEETRYGA 421
 DB 346 ADLPLPLLIOLSSVSGPLDQVAAASVGNRVEMMLRKYARLQEBVAPNRREREYEPYKGA 405
 QY 422 IYLEPLRGHENIAVLDDSSMYNIMIKYVNGPDTLYRPGCKGCGECWEAPVKKRFR 481
 DB 406 IYLEPKPGMYEDVLVLDFFSMYINIMKYVLSPTYLEPGEPPDPGAVNAPEVGHFRFR 465
 QY 482 CPPEFPTVLERLLEKRVRAEMKKYPPDPSPRYLLIDEEOKLKYLANASYMGSGA 541
 DB 466 SPGEFVPOVLKSLVELKRAVEAKKYPDPSPPEFKLIDERQALRYMANAIYGLGVGA 525
 QY 542 RMYCRECAKAVTAMGRHLITAINIARKLGLKYIGDPTSLFTYDPEKENFIKIKEE 601
 DB 526 RMYKREVAESVTAFAAILLDVIEQARRLGIVVYGDPTSLFYKKGKD-VDKLIKVEEK 584
 QY 602 LGFEIKLEKVKYKLPFTEAKKRYAGLIEDGRIDIVGFEAVRGWCELAKEVQKVEIVL 661
 DB 585 YGIDIKVDKDYAKVLPFEAKKRYAGLIDGRIDIVGFEAVRGWCELAKEVQKVEIVL 644
 QY 662 KTSRVNKA-----VEYVRKIVKELEEGKVPTEKVIWTKLSKLEEYTTTEPHYVAAKRL 717
 DB 645 KSRDIVEARRHGVILKYIREILIERLKNYKFNIDDLIWTTLKLEDEKAYPPHVAAGILK 704
 QY 718 SAGRYVSPDGKIGVYVYVGGGRISQRPAMPYFMVKDPSQIDVTVYVHDQIIPALRIIGYF 777
 DB 705 RHGRVVGKGTIGVIVYKGEKYSERLPIYLLDDIKIDIDYIERQIIPALRIAEVI 764
 QY 778 GITEKKLKASATG--OKTLPFLA 799
 DB 765 GVXESDLK---TGRMERSLDFLS 785
 RESULT 9
 AAY52022 standard; Protein; 781 AA.
 XX
 XX AAY52022;
 AC
 XX
 DT 12-JUL-2000 (first entry)
 XX
 XX A. fulgidus Af0497 protein.
 XX
 DE
 KM Polymerase; thermostable; nucleic acid elongation; sliding clamp protein;
 KM amplification; reverse transcription.
 OS Archaeoglobus fulgidus.
 XX
 PN DE19840771-A1.
 PD 10-FEB-2000.
 XX
 PF 07-SEP-1998; 98DE-1040771.
 XX
 PR 06-AUG-1998; 98DE-1035653.
 XX
 PA (LION-) LION BIOSCIENCE AG.
 XX
 PI Voss H, Moeckel G, Kober I, Kilger C;
 XX
 DR MPI; 2000-207149/19.
 XX
 PT A thermostable in vitro polymerase complex for template-dependent
 PT elongation of nucleic acids in amplification or reverse transcription
 PT methods -
 PS
 PS Disclosure; Page 72-75; 152pp; German.
 XX
 CC This invention describes a novel thermostable in vitro complex for

CC template-dependent elongation of nucleic acids which comprises a
 CC thermostable sliding clamp protein, which is connected with an
 CC elongation protein that shows thermostable polymerase activity. The
 CC thermostable in vitro accessory complex can be used to produce the
 CC elongation of nucleic acids, e.g. for amplification or reverse
 CC transcription. This is useful for sequencing nucleic acids by the
 CC polymerase chain reaction or reverse transcriptase-polymerase chain
 CC reaction (RT-PCR). The complex can be used to mark nucleic acids.
 CC AAY52000-Y52084 and JAW9752-W9799 represent proteins and protein
 CC fragments used to illustrate the method of the invention.

XX Sequence 781 AA;

Query Match 37.8%; Score 1581; DB 21; Length 781;
 Best Local Similarity 42.5%; Pred. No. 1.5e-139;
 Matches 350; Conservative 133; Mismatches 273; Indels 68; Gaps 18;

QY 1 MTEVFTVLDSSYEVEGKEPOVLIWGIAENGERVVLIDRSFRPYFALLAPGADPKOVAQ 60
 1 MERVGWLIDADYETIGGKAVVLMCKDDQG-IFVAYDYNEDPYFYI---GVDEBILKN 56
 DB 61 RIRALSRPKSPIIGVEDDKRYGRRPRVLRIRTVLPEAVREYRELKQVNDGVEDYLEAD 120
 57 AATSTRREVILKSKFEKAQKLTGREGVEGYIVAHHPQHVPKLRDYLQSG--DYREAD 113
 QY 121 IRFARVYLIDHDLPPFTWVVEAEPLNKG-----PRVDVYLVKSRPEPLYGALAPT 175
 114 IPFARVYLIDHDLADMGIAIBE-----KGGVIRSKYIEV-----ERIRPM 157
 QY 176 KLPLRLIAPDIEVYSKQSPREPDVYVIAVTDGDEVYLFIAESKDRKRIPEVY 235
 158 EPEELKMLVDCENLSPGMPBEKDPILVISKINDDEILITG--DERKTIIDFPVL 214
 QY 236 VKRIPDPIIVGNNHFDWYLLRRARILGKLDVTRVGAEPPTSVH--GHVSVGRIN 293
 215 IKSYPDPIIVGNODAFDWPYLKRKAERNWIPLD---VGSDSNVVFGRGKPKITGRIN 270
 QY 234 VDLVYAEEMPEIKISLEBAEYLGWKKSERVYINWMEIPYWDPPKRPLLLQVARD 353
 211 VDLVYAMRISDIKIKLEVAEFLGT--KEIADIADKIYVWGRGKEK-VINYARQ 327
 QY 354 DVRAVYGLAEKILPEALISYVYGLPLQOVGASVGRLEWYILRAAFKKEILVFNVER 413
 328 DAINTYLIKELLPMHYELSKMIRLPVDVTRMGSKQVWMLLSAKKIGIETANPREH 387
 QY 414 PEETRGALVLEPLRGVHENIADVSSWYPMIKVNGVDPLVPRKCGEGCGMREAP 473
 388 -AESYEGAFVLEPERGHEHVACLDPAWVPSIMAFNISPDY-----GCRD-DCEAP 440
 474 EVKIRFRPCPPGPFKTVLELLELRKRVRAEMKKYPPDSPSEYLLDERQKALKVLANSY 533
 441 EVGKHFRKSPDGFPRKILMLLIEKRELKVELKSLSPSEYKLLDKQDTLAKVLTNSFY 500
 QY 534 GYMMSGARWYCRECAKATYANGRHILRTAINIARAKLGKVIYIGTDSLFV---YDEEK 590
 501 GYKGNLAKWYCHPCAEATTANGRFIRISAKIASWGFVYLGDDDSIFVTKAGMTXD 560
 QY 551 VENTIKIIEBELGEIKLEKVKRLPFTPAKKRYAGLLEDGRIDIVGFAVAGWDCELAK 650
 561 VDRILDKLHELPIQIEVDEYSALFVE-KRVAAGLTGEGRLVVVAGLEVRGDCBELAK 619
 QY 651 EVQTKVVEIYVLTSEVNAKAVEYVKIYKLEBEQKPIELKIVIKTLSKLEETEAHAV 710
 620 KYOREVIEVILKEKNPEKLSLVKQVILIRIKEGKSLSEBAVYIKGLTKPSKYSQMAHV 679
 DB 711 VAAARMLSAGRVSPDGKIGYVYVKGGRISORAMPFMYVD-----PSQ 755
 680 KALKAREMGIIYVSSKIGYVYVKGSGNIGRAVPIDLIEDBGENLRIKTKSGIETKK 739
 QY 756 IDVYVYDHOIIPALRIIGYFGITKKLXASATGQTLDFLA 799
 DB 740 LDKVYIINDQIITSVRIERROYTEASLKSS--QMSLDSFFS 781

RESULT 10
 ID AAY51651 standard; protein; 781 AA.

XX AAY51651;

XX 01-JUN-2000 (first entry)

XX A. fulgidus AF0497 protein fragment.

XX Thermostable; template-dependent elongation; staple protein;
 XX elongation protein; amplification; reverse transcription.

XX Archaeoglobus fulgidus.

XX WO200008164-A2.

XX 17-FEB-2000.

XX 06-AUG-1999; 99WO-DE02480.

XX 06-AUG-1998; 98DE-1035653.

XX 07-SEP-1998; 98DE-1040771.

XX 18-JUN-1999; 99EP-0111795.

XX (LION-) LION BIOSCIENCE AG.

XX Kilger C, Kober I, Voss H, Moeckel G;

XX WPI; 2000-199576/17.

XX Thermostable in vitro polymerase complex for template-dependent
 XX elongation of nucleic acids

XX Disclosure; Page 167-170; 233pp; German.

XX This invention describes a novel thermostable in vitro complex (I) for
 CC template-dependent elongation of nucleic acids (NA) comprising a
 CC thermostable 'staple' protein and an elongation protein. The thermostable
 CC in vitro complex is useful for template-dependent elongation of NA, e.g.
 CC for amplification or reverse transcription. This is useful for sequencing
 CC NA by the polymerase chain reaction or reverse transcriptase-PCR. The
 CC complex can be used to mark NA. This sequence represents the Archaeoglobus
 CC fulgidus AF0497 protein which is used to illustrate the method of the
 CC invention.

XX Sequence 781 AA;

Query Match 37.8%; Score 1581; DB 21; Length 781;
 Best Local Similarity 42.5%; Pred. No. 1.5e-139;
 Matches 350; Conservative 133; Mismatches 273; Indels 68; Gaps 18;

QY 1 MTEVFTVLDSSYEVEGKEPOVLIWGIAENGERVVLIDRSFRPYFALLAPGADPKOVAQ 60
 1 MERVGWLIDADYETIGGKAVVLMCKDDQG-IFVAYDYNEDPYFYI---GVDEBILKN 56
 DB 61 RIRALSRPKSPIIGVEDDKRYGRRPRVLRIRTVLPEAVREYRELKQVNDGVEDYLEAD 120
 57 AATSTRREVILKSKFEKAQKLTGREGVEGYIVAHHPQHVPKLRDYLQSG--DYREAD 113
 QY 121 IRFARVYLIDHDLPPFTWVVEAEPLNKG-----PRVDVYLVKSRPEPLYGALAPT 175
 114 IPFARVYLIDHDLADMGIAIBE-----KGGVIRSKYIEV-----ERIRPM 157
 QY 176 KLPLRLIAPDIEVYSKQSPREPDVYVIAVTDGDEVYLFIAESKDRKRIPEVY 235
 158 EPEELKMLVDCENLSPGMPBEKDPILVISKINDDEILITG--DERKTIIDFPVL 214
 QY 236 VKRIPDPIIVGNNHFDWYLLRRARILGKLDVTRVGAEPPTSVH--GHVSVGRIN 293
 215 IKSYPDPIIVGNODAFDWPYLKRKAERNWIPLD---VGSDSNVVFGRGKPKITGRIN 270

RESULT 12

AAE22114 ID AAE22114 standard; Protein; 768 AA.

AAE22114;

25-JUL-2002 (first entry)

Archaeoglobus lithotrophicus DNA polymerase.

DNA polymerase; thermostable; enzyme.

Archaeoglobus lithotrophicus.

MO200220735-A2.

14-MAR-2002.

06-SEP-2001; 2001WO-US28007.

06-SEP-2000; 2000US-0656309.

(DIVE-) DIVERSA CORP.

Callen W, Mathur EJ, Short JM;

WPI; 2002-362247/39.

N-PSDB: AAD35183.

New thermostable polymerase useful for sequencing DNA, amplifying double stranded DNA, or incorporating a non-natural nucleotide or a nucleotide analog into a DNA molecule

Claim 42; Page 139-141; 161pp; English.

The invention relates to thermostable DNA polymerases having high temperature polymerase activity, such as those derived from *Pyrobolus fumaria* and nucleic acid molecules encoding such polymerases. Polymerases are useful for catalysing the formation or repair of a nucleic acid sequence and for modifying small molecules. They are also useful for sequencing DNA molecules, for preparing cDNA from mRNA, for amplifying double stranded DNA molecules and for incorporating non-natural nucleotides or nucleotide analogues into a DNA molecule. The present sequence is *Archaeoglobus lithotrophicus* DNA polymerase.

Sequence 768 AA;

Query Match 37.3%; Score 1563; DB 23; Length 768;

Best Local Similarity 41.3%; Pred. No. 7.4e-138; Matches 338; Conservative 154; Mismatches 273; Indels 54; Gaps 18;

QY 1 MTEVFTVLDSSIEVVGKEPOVITWGIAENGERVVLIDRSPRPYFALIAPADPKOVAQ 60
 DB 1 MTKKGMWLDADYITENDRAVIRLWCKDEG-ITFAYDSFOFPFALKEBSITAEIVK 59
 QY 61 -RIPALSRPKSPIIGVEDDKRYFGPRPRVLIRTVLPEAVREYRELKXNVGVEDVLEA 119
 DB 60 IKVQTKKEVITP-LKVEETTAQNLGRVEVEFKIYARHPOHVPKLRREVSO--YLEIREA 115
 QY 120 DIRFAMRYLIDHDLFPFTYRYVAEPLBNKGPFVDKYLK--SRPEPLUGEALAPKL 177
 DB 116 DIPFAYRYLIDKMLACMDGVIEGVERREK-GLRCYEIKRIERDSHOD-----F 163
 QY 178 PDLRLAFDIEVVSQKSGSPREPDPIYAVATDGDGEVFLIAGDODKRPREFVEYVK 237
 DB 164 PELKVMAFDCEMLSVGMDFPEKDPPIVISIKSGYEIL---NGNRELLTRVVKIR 220
 QY 238 RYDDIIVGVNNHFPMPYLARRALIGKLDVRRGCAEPTSVGHVSGRANLTX 297
 DB 221 DIDDIIIVGNOSDFMPYIKKAEKLVKLDIGR--DRSELIRGGRPKIKGRANVLY 278
 QY 298 DYAEEMPEIKIKSLEVAEYLVGMKKSERVLIINMWEIPDYWDPKGRPILLOYARDVRA 357

DB 279 DIAMSLDVKKLEBNVAEEFG--KTEILDIAEADQIKTKMTS-GDRSEYIKSRDILH 335
 QY 358 TYGLAEKILPPALIQSLVYTGLPIDOVGAMSVGFRLEWYLIPAAFKKELVFNVERPEET 417
 DB 336 TYFLAEELPMHVELSRNIRIPDDVTRSGRQKQVEMLLSEAHKDELAPNREM-AUS 394
 QY 418 YRGALVLEPLRGVHENIAVLDFSSMYPNIMIKYVGPDTLVRPEKCGECGEAEAEVQ 477
 DB 395 YEGAFVLEPARGLHENVICLDFASMPPSIMISYNISPTLVY--GKDDCNV--AEVEGH 450
 QY 478 RFRCPGPFKTVLERLEBKRVRAEMKKYPPDSPEYRLDEROKALKYLANASYGYNG 537
 DB 451 KFRKHDPGPFKRIIMLIEKRREIKKVMCTLDYNSPEYKLDIKQATLKVLNTSPGYTG 510
 QY 538 WSGARWYCEGCAKAVTAMGRHLIRTAIINIRKLGKVIYGDTSLPVYTPPEKVENFIKI 597
 DB 511 WSLARWYCEGCBATAMGRHPIKTSARIKELGFEVLVYGDTSIFVKKDGSLSEELKE 570
 QY 598 IK-----EELGFEIKLEKVRKRLPFTAKKRVAGLLEDRIDIVGEFAVRGDMCELAK 650
 DB 571 VKTLQKLSSEMPQIIEIDEYETIFVE-KRVAGLQDORILIVKGLVLRGDMCELAK 629
 QY 651 EVQTKVVEIVLKTSEVNKAVEYVRKIVKLEBGKPIRKLYIMKTLKSLSEYTTAPVY 710
 DB 630 KIQKGVIEIILEKQNPKEPAABVGVIEIKKQKLPLEDYIITVQGLTRKPSKYESMQAHV 689
 QY 711 VAAKRLMSAGYRVSPGDKIGYIVKGGRIISORAMPYMYKD-----PSQIDV 758
 DB 690 KAAMKAKRGVIVYITGSKVGFVTKGVNIGRAPSDILDFGEVITDLDGNKKYIKD 749
 QY 759 TYYVDHQIIPALNRIYFGITTEKKLKASATQKTLPDF 797
 DB 750 EYVIDHQVLPVLRILSRFGYTEALQKGAEE-QQTLDAF 787

RESULT 13
 AAE13416 ID AAE13416 standard; Protein; 774 AA.
 XX AAE13416;
 AC AAE13416;
 DT 12-FEB-2002 (first entry)
 XX P. kodakarensis KOD1 strain mutant DNA polymerase, H147K.
 XX DNA polymerase; thermostable; 3'-5'-exonuclease; improved fidelity;
 XX amplification; extension; mutation; mutant; mutain.
 OS *Pyrococcus kodakarensis*.
 XX Synthetic.
 FH Key Location/Qualifiers
 FT Msc-difference 147
 FT note="Wild-type His substituted by Lys"
 XX EPI154017-A1.
 XX 14-NOV-2001.
 XX 08-MAY-2001; 2001EP-0111032.
 XX 11-MAY-2000; 2000JP-0138796.
 XX (TOWN) TOYO BOSEKI KK.
 XX Kuroita T, Kitabayashi M, Ishida Y, Komatsubara H, Nishiya Y,
 PI Kawasumi B, Kawamura Y, Imanaka T;
 XX WPI; 2002-012658/02.
 XX Modified heat stable polymerase, useful in polymerase chain reactions,
 PT has specific histidine residue mutated to impart e.g. increased


```

QY 184 AFDIEVYSKQSPRPERDPVIVIAVKTDDGDEVL-----FLAEKDDRKPIREFVEY 235
DB 139 AFDIEVYSKQSPRPERDPVIVIAVKTDDGDEVL-----FLAEKDDRKPIREFVEY 197
QY 236 VKRYDPDIIIVGNNHFPWYLLRRARIGIKLDVTRVGAEPPTSVGH---VSYPGR 292
DB 198 IREKDPDIIIVTNGDSFDFPLAKBAKGIKLTIGRD-GSEPKQKQIGMTAVEXKGI 256
QY 293 NVDLYDAEMPEIKIKSLERVAEYLGWKKSERVITIMWBEIPYMDPKRLLQYAR 352
DB 257 HDPLVIVITRTINLPYTLAVVEAL-FGKPEKRYAD-BIAAMSGENLERVAKISM 313
QY 353 DVPRTATYGLAEKILPFAIQLSYVTGLPDOVGAMSVGRLEWYLIRAFPMKELVNRVE 412
DB 314 EDKATYELGKEFLPMEIQLSRVGOPLMDVSRSSSTGNLVWFLIRAYERNVADNPKS 373
QY 413 RPE-----ETVGAIVLEPLRGVHENIAVLDFSSMYPNIMIKNVGPDTLVPRGKCGE 466
DB 374 EEEYQRRLSESYTGGFVKEPEKIMENIVYLDPRALYPSIITTHVSPDTLANEG----- 428
QY 467 CGCWE-APEVKARFRRCPPGFVTKVLERLLELRKVRAMKCPDPSPEYRLDDEKQAL 525
DB 429 CKRYDIAPOVGHKFCOIPGPIPSLGLHLEEROKIKTKME-TODPIEKILLDYRQKAI 487
QY 526 KYLANASGYMGSGARWYCRECAKAVTAMGRHLIRTA-INIARKGLKVIYGDPTSLFV 584
DB 488 KILANSFYGYGVARARWYCKECSASTAMGRKYLEIWKLEERKFGKVLVYIDTGLVA 547
QY 585 TYD-----PEKVENPKIKIEELG--PEIKLEKVVYRLPFTAKRYAGLLEDRIDI 635
DB 548 TIRGSESEIKKALEFKYKINSKLPGLLELEEGTKRGFFV-TKGRVAVIDEKRAVIT 606
QY 636 VGEAVRGMCELAKEVOTKVEIVLKTSEVNKAVEYRKIVKLESGKVPLEKVIYMT 695
DB 607 RGLLEIVRDMSEIAKEVOTARLETILHGVGEAVARIVAEVIAQKLANVELIPPEKLAIVQ 666
QY 696 LSKRLSEYTTFAPHVAAKRLMISAGRVSPGDKIGVYIVKGGRIISORAMP---YEMVMD 752
DB 667 ITRPLHEYKAIQPHVAAKRLMISAGRVSPGDKIGVYIVKGGRIISORAMP---YEMVMD 722
QY 753 PS--QIDVYVYVDHIIIPALRLIGVIGITEKKKASATGCKTFDFPL-AKKS 802
DB 723 PKGKYDAEYIENQVLAVALRILEGFYRKEDLRYQTRQVGLTSMILNKKS 775

```

RESULT 15

AAE13417 standard; Protein: 774 AA.

AAE13417;

12-FEB-2002 (first entry)

P. kodakarensis KOD1 strain mutant DNA polymerase, H147R.

DNA polymerase; thermostable; 3'-5'-exonuclease; improved fidelity;

amplification; extension; mutation; mutant; mutain.

Pyrococcus kodakarensis.

Synthetic.

Key

FT Msc-difference 147

PN EPI154017-A1.

PD 14-NOV-2001.

PR 08-MAY-2001; 2001EP-0111032.

PR 11-MAY-2000; 2000JP-0138796.

```

XX (TOYM) TOYO BSEKI KK.
PA Kurokita T, Kikawaseh M, Ishida Y, Komatsubara H, Nishiya Y,
PI Kawakami B, Kawamura Y, Imanaka T;
PI WPI, 2002-012658/02.
DB Modified heat stable polymerase, useful in polymerase chain reactions,
XX has specific histidine residue mutated to impart e.g. increased
XX fidelity.
XX Claim 12, Page - 64pp; English.
XX The invention relates to modified thermostable DNA polymerase
XX having 3'-5'-exonuclease activity from Pyrococcus kodakarensis.
XX In the modified thermostable DNA polymerase, a specific histidine
XX residue in the wild-type enzyme has been replaced by some other
XX amino acid. The modified thermostable DNA polymerase has enhanced
XX amplification efficiency and/or improved fidelity in PCR reaction
XX and is used as reagent for amplification, extension or mutation of
XX template DNA in essentially standard polymerase chain reactions.
XX The present sequence is Pyrococcus kodakarensis KOD1 strain
XX mutant DNA polymerase.
XX Note: This sequence is not shown in the specification, but is
XX derived from the thermostable DNA polymerase (AAE13405) shown in
XX page 36-42 of the specification.
XX Sequence 774 AA:
XX
XX Query Match 29.7%; Score 1245.5; DB 23; Length 774;
XX Basic Local Similarity 36.7%; Pred.No. 6.4e-108;
XX Matches 304; Conservative 145; Mismatches 287; Indels 93; Gaps 22;

```

```

QY 8 VLDSSYEVGKPEQVIINGIAENGERVVLIDSPFPYFALLAGADPKV---AQRIR 63
DB 2 ILDTIDYITDGDGPVIRIK-KENGEFKIEYDRFEPYFALLKDSAIIEVKKITERRG 60
QY 64 ALSRPSPIIGVEDDKRYFGPRPVLRIRTVLPEAVEYRELWGNVDEVDLEADIRF 123
DB 61 TVVTYGR-----VEKQKFLGRPVEWKLFTPHQVPAIRDKIRBPVAVIDYEIDIF 116
QY 124 AMRYLLDHDLPFTVYRVAEPLLENKQGFVADVLSRREPILYGRALAPTLPLRL 183
DB 117 AKRYLIDKGLV-----PMEGD-----BELKRL 138
QY 184 AFDIEVYSKQSPRPERDPVIVIAVKTDDGDEVL-----FLAEKDDRKPIREFVEY 235
DB 139 AFDIEVYSKQSPRPERDPVIVIAVKTDDGDEVL-----FLAEKDDRKPIREFVEY 197
QY 236 VKRYDPDIIIVGNNHFPWYLLRRARIGIKLDVTRVGAEPPTSVGH---VSYPGR 292
DB 198 IREKDPDIIIVTNGDSFDFPLAKBAKGIKLTIGRD-GSEPKQKQIGMTAVEXKGI 256
QY 293 NVDLYDAEMPEIKIKSLERVAEYLGWKKSERVITIMWBEIPYMDPKRLLQYAR 352
DB 257 HDPLVIVITRTINLPYTLAVVEAL-FGKPEKRYAD-BIAAMSGENLERVAKISM 313
QY 353 DVPRTATYGLAEKILPFAIQLSYVTGLPDOVGAMSVGRLEWYLIRAFPMKELVNRVE 412
DB 314 EDKATYELGKEFLPMEIQLSRVGOPLMDVSRSSSTGNLVWFLIRAYERNVADNPKS 373
QY 413 RPE-----ETVGAIVLEPLRGVHENIAVLDFSSMYPNIMIKNVGPDTLVPRGKCGE 466
DB 374 EEEYQRRLSESYTGGFVKEPEKIMENIVYLDPRALYPSIITTHVSPDTLANEG----- 428
QY 467 CGCWE-APEVKARFRRCPPGFVTKVLERLLELRKVRAMKCPDPSPEYRLDDEKQAL 524
DB 429 CKRYDIAPOVGHKFCOIPGPIPSLGLHLEEROKIKTKME-TODPIEKILLDYRQKAI 485
QY 526 KYLANASGYMGSGARWYCRECAKAVTAMGRHLIRTA-INIARKGLKVIYGDPTSLFV 583
DB 488 KILANSFYGYGVARARWYCKECSASTAMGRKYLEIWKLEERKFGKVLVYIDTGLVA 545

```


Qy 584 VTY---DPEKVE---NFIKIIKEEL--GPEIKLEKXYKRLFPTTEAKKRYAGLLEDGRID 634
Db 546 ATIPGADAEIVYKKAMEFLKTIINAKLPGALELEYEGFYKGFV-TKKYIAVIDEKGKIT 604
Qy 635 IVGEFAVRGDMCELAKEVQTVVEIVLKTSEVNNAVEYVRKIVKELEBEGKPIEKLYWK 694
Db 605 TRGLEIVRRDMSIASETOARVLEALLKGDYKAVRIKVEVTEKLSKYEVPEKLYIHE 664
Qy 695 TLSKRLBEYTEAPHVVAAKRMLSNQYRVSFGDKIGYIVKGGRIISORANPYFMVKDPS 754
Db 665 QITRDLKDYKATGPHVAVAKRLAARGVIRPGTVISYIVLKGSGRIGDRAIP-FDEFDPT 723
Qy 755 --QIDVTYVVDHOLIIPALRILGYFGITEKKLKASATGOKTLFDFLANK 801
Db 724 KHKYDAEYIENQVLPVAVERTILRAFGYRKEDLRYOKTRQVGLSAMLKPK 772

arch completed: April 23, 2003, 07:30:23
p time : 46 secs

10

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:30:27 ; Search time 252 Seconds
(without alignments)
255.334 Million cell updates/sec

Title: US-10-034-849-2

Perfect score: 4188
Sequence: 1 MTEVFTVLDSSEYVWGKEP.....LKASATGQKTLFPLAKSK 803

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB pep:*
- 2: /cgn2_6/prodata/2/pubppa/PCR_NEW_PUB pep:*
- 3: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB pep:*
- 4: /cgn2_6/prodata/2/pubppa/US06_PUBCOMB pep:*
- 5: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB pep:*
- 6: /cgn2_6/prodata/2/pubppa/US07_PUBCOMB pep:*
- 7: /cgn2_6/prodata/2/pubppa/PCR_PUBCOMB pep:*
- 8: /cgn2_6/prodata/2/pubppa/US08_PUBCOMB pep:*
- 9: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB pep:*
- 10: /cgn2_6/prodata/2/pubppa/US09_PUBCOMB pep:*
- 11: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB pep:*
- 12: /cgn2_6/prodata/2/pubppa/US10_PUBCOMB pep:*
- 13: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB pep:*
- 14: /cgn2_6/prodata/2/pubppa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4188	100.0	803	9	US-10-029-382-2
2	4188	100.0	803	10	US-09-391-340-4
3	4188	100.0	803	10	US-09-948-369-4
4	4188	100.0	803	10	US-09-948-369-14
5	4188	100.0	803	12	US-10-034-849-2
6	4188	100.0	803	12	US-10-034-849-2
7	4164	99.4	803	10	US-09-948-369-16
8	1563	37.3	788	10	US-09-391-340-6
9	1563	37.3	788	10	US-09-948-369-6
10	1242.5	29.7	774	10	US-09-852-922-2
11	1194	28.5	774	10	US-09-803-165-34
12	564.5	13.5	844	9	US-10-027-806-30
13	564.5	13.5	844	9	US-10-034-623-30
14	564.5	13.5	844	9	US-10-027-801-30
15	552	13.2	845	9	US-10-027-806-62
16	552	13.2	845	9	US-10-034-623-62
17	552	13.2	845	9	US-10-027-801-62
18	542.5	13.0	877	10	US-09-391-340-8
19	542.5	13.0	877	10	US-09-948-369-8

20	506.5	12.1	762	10	US-09-391-340-10	Sequence 10, Appl
21	506.5	12.1	762	10	US-09-948-369-10	Sequence 10, Appl
22	452	10.8	1235	10	US-09-904-065-17	Sequence 17, Appl
23	452	10.8	1235	10	US-09-904-065-4	Sequence 4, Appl
24	452	10.8	1240	10	US-09-904-065-15	Sequence 15, Appl
25	450	10.7	1235	10	US-09-904-065-8	Sequence 8, Appl
26	450	10.7	1235	10	US-09-904-065-10	Sequence 10, Appl
27	450	10.7	1235	10	US-09-904-065-18	Sequence 18, Appl
28	450	10.7	1235	10	US-09-904-065-19	Sequence 19, Appl
29	449	10.7	1235	10	US-09-904-065-6	Sequence 6, Appl
30	449	10.7	1235	10	US-09-904-065-16	Sequence 16, Appl
31	445	10.6	1238	10	US-09-904-065-2	Sequence 2, Appl
32	445	10.6	1238	10	US-09-904-065-14	Sequence 14, Appl
33	368	8.8	1242	10	US-09-904-065-12	Sequence 12, Appl
34	368	8.8	1242	10	US-09-904-065-13	Sequence 13, Appl
35	167	4.0	327	10	US-09-925-300-1209	Sequence 1209, Ap
36	113.5	2.7	517	12	US-10-062-254-328	Sequence 328, App
37	112.5	2.7	159	9	US-10-055-364-50	Sequence 50, App
38	112.5	2.7	405	9	US-10-102-806-445	Sequence 445, App
39	111.5	2.7	599	10	US-09-815-242-11326	Sequence 11326, A
40	109	2.6	833	9	US-09-940-9258-85	Sequence 85, Appl
41	109	2.6	1211	10	US-09-815-242-11608	Sequence 11608, A
42	106.5	2.5	550	10	US-09-815-242-10499	Sequence 10499, A
43	106	2.5	893	9	US-09-891-3328-9	Sequence 9, Appl
44	105.5	2.5	693	10	US-09-815-242-13233	Sequence 13233, A
45	104.5	2.5	777	10	US-09-801-574-88	Sequence 88, Appl

ALIGNMENTS

RESULT 1
US-10-029-382-2
Sequence 2, Application US/10029382
Patent No. US20020164618A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: CALLEN, Walter
APPLICANT: MATHER, Eric
TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS
FILE REFERENCE: DIVER1350-2
CURRENT APPLICATION NUMBER: US/10/029,382
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: US/09/656,309
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 803
TYPE: PRT
ORGANISM: Pyrobolus fumarius
US-10-029-382-2

Query Match 100.0%; Score 4188; DB 9; Length 803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEVFTVLDSSEYVWGKEPQVIIMGIAENGERVLLDRSFRPYFVLLAPGADPKOVAQ 60
Db 1 MTEVFTVLDSSEYVWGKEPQVIIMGIAENGERVLLDRSFRPYFVLLAPGADPKOVAQ 60

Qy 61 RIRALSPKSPITIGVEDDKRRYFGRPRVLRIRTVLPEAVREYRELKNDVGVEDVLEAD 120
Db 61 RIRALSPKSPITIGVEDDKRRYFGRPRVLRIRTVLPEAVREYRELKNDVGVEDVLEAD 120

Qy 121 IRPFRVLLIDHDLPEFWYVEABPLENKGPRVDKYLKSPREPLVGGALPTKLPDL 180
Db 121 IRPFRVLLIDHDLPEFWYVEABPLENKGPRVDKYLKSPREPLVGGALPTKLPDL 180

Qy 181 RILAFDIEVSKQSPPERPDPVIVIAVKTDDGDEVLFIAEGKDRKPIEFVYVRYD 240
Db 181 RILAFDIEVSKQSPPERPDPVIVIAVKTDDGDEVLFIAEGKDRKPIEFVYVRYD 240

Wed Apr 23 08:09:12 2003

us-10-034-849-2.rapb

Page 2

Qy 241 PDIIIVGNNHFPMPYLLRRARILGILKDVTRVGAEPPTTSVHGHSVFGRLNVDLYDA 300
Db 241 PDIIIVGNNHFPMPYLLRRARILGILKDVTRVGAEPPTTSVHGHSVFGRLNVDLYDA 300
Qy 301 EEMPEIKISLEEVAEYLGWKKSERVIINMMEIPDYWDDPKRPLLOQYARDVATYG 360
Db 301 EEMPEIKISLEEVAEYLGWKKSERVIINMMEIPDYWDDPKRPLLOQYARDVATYG 360
Qy 361 LAEKILPFAIQLSYVTGLPLDOYGAMSVGFRLIEWLIRAAFKMKELVNNRVERPEETYG 420
Db 361 LAEKILPFAIQLSYVTGLPLDOYGAMSVGFRLIEWLIRAAFKMKELVNNRVERPEETYG 420
Qy 421 AIVLEPLRGVHENIAVLDPSSMYPNIMIKNVGPDTLVRPGEKCGEGCEAEPEVGRFR 480
Db 421 AIVLEPLRGVHENIAVLDPSSMYPNIMIKNVGPDTLVRPGEKCGEGCEAEPEVGRFR 480
Qy 481 RCPGPFKTVLERLELRLKRVRAEMKKYPPDSPERYLLDEROKALKVLANASGYMGWG 540
Db 481 RCPGPFKTVLERLELRLKRVRAEMKKYPPDSPERYLLDEROKALKVLANASGYMGWG 540
Qy 541 ARMYCBECAKAVTAMGSHLIRTAINTARKLGKVIYGTDSLFTYDPEKVENFIKIIE 600
Db 541 ARMYCBECAKAVTAMGSHLIRTAINTARKLGKVIYGTDSLFTYDPEKVENFIKIIE 600
Qy 601 ELGFEIKLEKYKRLPFTTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVOTKVEIV 660
Db 601 ELGFEIKLEKYKRLPFTTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVOTKVEIV 660
Qy 661 LKTSVNRKAVEYVRKIVKELEEGKVPLEKLVIMKTLSEYTTTEAPHVAAKRLSAG 720
Db 661 LKTSVNRKAVEYVRKIVKELEEGKVPLEKLVIMKTLSEYTTTEAPHVAAKRLSAG 720
Qy 721 YRSPGDKIGVYVKGGRISORAMPYFMVXDPQSDIVTYVVDHQIIPALRLILGFGIT 780
Db 721 YRSPGDKIGVYVKGGRISORAMPYFMVXDPQSDIVTYVVDHQIIPALRLILGFGIT 780
Qy 781 EKKLKASATGQKTLFDFLAKSK 803
Db 781 EKKLKASATGQKTLFDFLAKSK 803

RESULT 2
US-09-391-340-4
; Sequence 4, Application US/09391340A
; Patent No. US20020013455A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Mather, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/09/391,340A
; EARLIER FILING DATE: 1999-09-07
; EARLIER APPLICATION NUMBER: US 08/907,166
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 803
; TYPE: FRT
; ORGANISM: Pyrolobus fumarius
US-09-391-340-4

Query Match 100.0%; Score 4188; DB 10; Length 803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEVFTVLDSSYEYVKGKPOYIINGIAENGERVLLIDRSFRPYFALLAPADKQYAO 60
Db 1 MTEVFTVLDSSYEYVKGKPOYIINGIAENGERVLLIDRSFRPYFALLAPADKQYAO 60
Qy 61 RIRALSRPKSPIIGVEDDKRYKFGRRVLAIRTVLEAVREVELVKNVDGEVLEAD 120
Db 61 RIRALSRPKSPIIGVEDDKRYKFGRRVLAIRTVLEAVREVELVKNVDGEVLEAD 120

Db 61 RIRALSRPKSPIIGVEDDKRYKFGRRVLAIRTVLEAVREVELVKNVDGEVLEAD 120
Qy 121 IRPAMVYLIDHDLPEFTMYVEAEAPLENKMGGRVKNVLYKSRREPLYGALPTGLPDL 180
Db 121 IRPAMVYLIDHDLPEFTMYVEAEAPLENKMGGRVKNVLYKSRREPLYGALPTGLPDL 180
Qy 181 RIIAPDIEYYSKQSPREPERDPIVIAVKTDDGEVLFIAEKODRPIEFPEYKRYD 240
Db 181 RIIAPDIEYYSKQSPREPERDPIVIAVKTDDGEVLFIAEKODRPIEFPEYKRYD 240
Qy 241 PDIIIVGNNHFPMPYLLRRARILGILKDVTRVGAEPPTTSVHGHSVFGRLNVDLYDA 300
Db 241 PDIIIVGNNHFPMPYLLRRARILGILKDVTRVGAEPPTTSVHGHSVFGRLNVDLYDA 300
Qy 301 EEMPEIKISLEEVAEYLGWKKSERVIINMMEIPDYWDDPKRPLLOQYARDVATYG 360
Db 301 EEMPEIKISLEEVAEYLGWKKSERVIINMMEIPDYWDDPKRPLLOQYARDVATYG 360
Qy 361 LAEKILPFAIQLSYVTGLPLDOYGAMSVGFRLIEWLIRAAFKMKELVNNRVERPEETYG 420
Db 361 LAEKILPFAIQLSYVTGLPLDOYGAMSVGFRLIEWLIRAAFKMKELVNNRVERPEETYG 420
Qy 421 AIVLEPLRGVHENIAVLDPSSMYPNIMIKNVGPDTLVRPGEKCGEGCEAEPEVGRFR 480
Db 421 AIVLEPLRGVHENIAVLDPSSMYPNIMIKNVGPDTLVRPGEKCGEGCEAEPEVGRFR 480
Qy 481 RCPGPFKTVLERLELRLKRVRAEMKKYPPDSPERYLLDEROKALKVLANASGYMGWG 540
Db 481 RCPGPFKTVLERLELRLKRVRAEMKKYPPDSPERYLLDEROKALKVLANASGYMGWG 540
Qy 541 ARMYCBECAKAVTAMGSHLIRTAINTARKLGKVIYGTDSLFTYDPEKVENFIKIIE 600
Db 541 ARMYCBECAKAVTAMGSHLIRTAINTARKLGKVIYGTDSLFTYDPEKVENFIKIIE 600
Qy 601 ELGFEIKLEKYKRLPFTTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVOTKVEIV 660
Db 601 ELGFEIKLEKYKRLPFTTEAKKRYAGLLEDGRIDIVGFEAVRGDMCELAKEVOTKVEIV 660
Qy 661 LKTSVNRKAVEYVRKIVKELEEGKVPLEKLVIMKTLSEYTTTEAPHVAAKRLSAG 720
Db 661 LKTSVNRKAVEYVRKIVKELEEGKVPLEKLVIMKTLSEYTTTEAPHVAAKRLSAG 720
Qy 721 YRSPGDKIGVYVKGGRISORAMPYFMVXDPQSDIVTYVVDHQIIPALRLILGFGIT 780
Db 721 YRSPGDKIGVYVKGGRISORAMPYFMVXDPQSDIVTYVVDHQIIPALRLILGFGIT 780
Qy 781 EKKLKASATGQKTLFDFLAKSK 803
Db 781 EKKLKASATGQKTLFDFLAKSK 803

RESULT 3
US-09-948-369-4
; Sequence 4, Application US/09948369
; Patent No. US2002013223A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: MATHER, ERIC
; APPLICANT: CALLEN, WALTER
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS
; FILE REFERENCE: DIVER1350-3
; CURRENT APPLICATION NUMBER: US/09/948,369
; EARLIER FILING DATE: 2001-09-06
; EARLIER APPLICATION NUMBER: US 09/556,309
; EARLIER FILING DATE: 2000-09-06
; EARLIER APPLICATION NUMBER: US 09/391,340
; EARLIER FILING DATE: 1999-09-07
; EARLIER APPLICATION NUMBER: US 08/907,166
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4

```

; LENGTH: 803
; TYPE: PRT
; ORGANISM: Pyrolobus fumarius
US-09-948-369-4

```

Query Match	100.0%;	Score 4188;	DB 10;	Length 803;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 803; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	MTEVAFVLJDSSEYVGEKPEOVIIIMGLAENGERVLLIDRSRPRVFAVALLA PGADPKVOA	60
Dp	1	MTEVAFVLJDSSEYVGEKPEOVIIIMGLAENGERVLLIDRSRPRVFAVALLA PGADPKVOA	60
Qy	61	RIRALSREKSPBIIIGVEDDKRYKFGPRPRVLRIRVLPEAVREYELVKNVDGVEDVLEAD	120
Dp	61	RIRALSREKSPBIIIGVEDDKRYKFGPRPRVLRIRVLPEAVREYELVKNVDGVEDVLEAD	120
Qy	121	IRFAMRYLIDHDLPEFTYRVRVEABPLENKMGRFVDDKTYLVKSRPEPLYGEALAPTKLPDL	180
Dp	121	IRFAMRYLIDHDLPEFTYRVRVEABPLENKMGRFVDDKTYLVKSRPEPLYGEALAPTKLPDL	180
Qy	181	RILAFDIEVYSKQSGPRPERDPVIVIAVKTDDGDEVLFIAGXDRKPIREFAVEYKRYD	240
Dp	181	RILAFDIEVYSKQSGPRPERDPVIVIAVKTDDGDEVLFIAGXDRKPIREFAVEYKRYD	240
Qy	241	PDIIVGYNNNHPFDMPYLLRRARIIIGIKLDVTRRGABEPTTSHGHVSVPGRLANDLYDA	300
Dp	241	PDIIVGYNNNHPFDMPYLLRRARIIIGIKLDVTRRGABEPTTSHGHVSVPGRLANDLYDA	300
Qy	301	EEMPIKIKSLIEVAEYLGWKKSSRVINMMELPDYWDPKKRPILLOVARDVRYATYG	360
Dp	301	EEMPIKIKSLIEVAEYLGWKKSSRVINMMELPDYWDPKKRPILLOVARDVRYATYG	360
Qy	361	LAEKLPFAIOLSYVTGLPLLOVGAMSGVFLBMYLIRAPKMEELVNNRVERPEETRYG	420
Dp	361	LAEKLPFAIOLSYVTGLPLLOVGAMSGVFLBMYLIRAPKMEELVNNRVERPEETRYG	420
Qy	421	AIIVLEPLRGVHENIAVLDFFSSMYPNIMIKYVGPDTLVRPGEKCGECGWEAPEVKHFR	480
Dp	421	AIIVLEPLRGVHENIAVLDFFSSMYPNIMIKYVGPDTLVRPGEKCGECGWEAPEVKHFR	480
Qy	481	RCPPGFPTVLIERLLELKRVRAEKKKVPBPSPEYRLIDEBQKLYLANASVGYMWSG	540
Dp	481	RCPPGFPTVLIERLLELKRVRAEKKKVPBPSPEYRLIDEBQKLYLANASVGYMWSG	540
Qy	541	ARMYCRECAKAVTMGRHLIRTAIINIAKGLKATVYGTDSLPTYYPBEKVENFIKIIKE	600
Dp	541	ARMYCRECAKAVTMGRHLIRTAIINIAKGLKATVYGTDSLPTYYPBEKVENFIKIIKE	600
Qy	601	ELGFPITKLEKTYKRLFTEBAKKRVAAGLLEDRIDIVGEFAVRGDMCELANEVQTKVEIV	660
Dp	601	ELGFPITKLEKTYKRLFTEBAKKRVAAGLLEDRIDIVGEFAVRGDMCELANEVQTKVEIV	660
Qy	661	LKTSVNNNAVEVVRKIIVELBEGKVPLEKVIWMTLSRLESEYTEA PHVVAARMKSAG	720
Dp	661	LKTSVNNNAVEVVRKIIVELBEGKVPLEKVIWMTLSRLESEYTEA PHVVAARMKSAG	720
Qy	721	YRVSPGDKIGVIVYKVGGRISQRAMPYFMVNDPSQIDVTYVVDHQIIPALRLILGYGIT	780
Dp	721	YRVSPGDKIGVIVYKVGGRISQRAMPYFMVNDPSQIDVTYVVDHQIIPALRLILGYGIT	780
Qy	781	EKKUKASATGQKTLFDFLAKSK 803	
Dp	781	EKKUKASATGQKTLFDFLAKSK 803	

RESULT 4
US-09-948-369-14

```

; Sequence 14, Application US/09948365
; Patent No. US20020132243A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter

```

```

1  APPLICANT: MATTHEW, Eric
2  TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS
3  TITLE OF INVENTION: THEREOF
4  FILE REFERENCE: DIVER1350-3
5  CURRENT APPLICATION NUMBER: US/09/948,369
6  CURRENT FILING DATE: 2001-09-06
7  PRIOR APPLICATION NUMBER: US 09/656,309
8  PRIOR FILING DATE: 2000-09-06
9  PRIOR APPLICATION NUMBER: US 09/391,340
10 PRIOR FILING DATE: 1999-09-07
11 PRIOR APPLICATION NUMBER: US 08/907,166
12 PRIOR FILING DATE: 1997-08-06
13 NUMBER OF SEQ ID NOS: 16
14 SOFTWARE: PatentIn version 3.0
15 SEQ ID NO 14
16 LENGTH: 803
17 TYPE: prt
18 ORGANISM: Pyrololobus fumarius
19 US-09-948-369-14

```

Query Match	100.0%	Score 4188;	DB 10;	Length 803;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 803; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	MTEVAFVLDSSZYVWGPEPOVJIMGJLAENGERVLLIDRSFRPFVYLLAPGADPKVOAQ	60
Dd	1	MTEVAFVLDSSZYVWGPEPOVJIMGJLAENGERVLLIDRSFRPFVYLLAPGADPKVOAQ	60
Qy	61	RIRALSREKSPDIIGVEDDKRYCFGRPRPRVLIRITVLPEAVREYRELKXNDGVEDVLEAD	120
Dd	61	RIRALSREKSPDIIGVEDDKRYCFGRPRPRVLIRITVLPEAVREYRELKXNDGVEDVLEAD	120
Qy	121	IRFAMRYLIDHDLPFFMYRVEABPLNKGFRVDKYYLVKSREPELYGBALPTKLPDL	180
Dd	121	IRFAMRYLIDHDLPFFMYRVEABPLNKGFRVDKYYLVKSREPELYGBALPTKLPDL	180
Qy	181	RILAFDIEVYSKOSPPRPRPVIVIAVKTDDGDEVLFIAEGKODRKPIIEFVEYKRYD	240
Dd	181	RILAFDIEVYSKOSPPRPRPVIVIAVKTDDGDEVLFIAEGKODRKPIIEFVEYKRYD	240
Qy	241	PDIIVGYNNHFDMPYLLRRARILIGIKLDYTRRGAEPPTS5VGHVSVPRLNVDLDYA	300
Dd	241	PDIIVGYNNHFDMPYLLRRARILIGIKLDYTRRGAEPPTS5VGHVSVPRLNVDLDYA	300
Qy	301	EEMPEIKIKSLEEVAEYLVGWKKSERYIINMWEIPDYMDPKKRPLLLQYARDVRAATY	360
Dd	301	EEMPEIKIKSLEEVAEYLVGWKKSERYIINMWEIPDYMDPKKRPLLLQYARDVRAATY	360
Qy	361	LAEKILPPIALDSTVYTGPLDQVAMSVGRLEMYLIRAAFKMKELYPNRVERPEETRYG	420
Dd	361	LAEKILPPIALDSTVYTGPLDQVAMSVGRLEMYLIRAAFKMKELYPNRVERPEETRYG	420
Qy	421	AIYIEPLRGVHEINAVLDFSSMYNINIMKNVSGPDTLYVRGEGKGECCWEAPVVKHFR	480
Dd	421	AIYIEPLRGVHEINAVLDFSSMYNINIMKNVSGPDTLYVRGEGKGECCWEAPVVKHFR	480
Qy	481	RCPDGFECTVLERLLELRKRVRAMKCYPPDPSPEYRLDLDEROKALKYLAASGYMWSG	540
Dd	481	RCPDGFECTVLERLLELRKRVRAMKCYPPDPSPEYRLDLDEROKALKYLAASGYMWSG	540
Qy	541	ARWTCRECAKAVTMAGHRLIRTAINIRKGLKVIYGDTSLSFYTYDPEKEVNIKIKE	600
Dd	541	ARWTCRECAKAVTMAGHRLIRTAINIRKGLKVIYGDTSLSFYTYDPEKEVNIKIKE	600
Qy	601	ELGPEIKLEKYKRLFPTEAKKRVAGLLEGRIDIVGFPAVRGOWGLAEXVOTKVEIV	660
Dd	601	ELGPEIKLEKYKRLFPTEAKKRVAGLLEGRIDIVGFPAVRGOWGLAEXVOTKVEIV	660
Qy	661	LKTSSEVNAVAYYAKIKVLELEEGCAPIEKLVIWKTLSKRLSEYVTEBPHVAAKRMLSAG	720
Dd	661	LKTSSEVNAVAYYAKIKVLELEEGCAPIEKLVIWKTLSKRLSEYVTEBPHVAAKRMLSAG	720
Qy	721	YRVSBGDKIGVIVKGGGRISQRAMPYFMWVDSQSDIVTYVDHQIIPALRILGYFGIT	780

Wed Apr 23 08:09:12 2003

us-10-034-849-2.rapb

Page 4

Db 721 YRVSPGDKIGYIVYKGGGRISORAMPYFMVKDPSQIDVTYVVDHQQIIPALRILGYFGIT 780
Qy 781 EKKLKASATGOKTLFDFLAKSK 803
Db 781 EKKLKASATGOKTLFDFLAKSK 803

RESULT 5

US-10-034-849-2
Sequence 2, Application US/10034849
Patent No. US2002015108A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: CALLEN, Walter
APPLICANT: MATHER, Eric
TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY
TITLE OF INVENTION: AND METHODS OF USE
FILE REFERENCE: DIVER1350-2
CURRENT APPLICATION NUMBER: US/10/034, 849
CURRENT FILING DATE: 2001-12-21
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 803
TYPE: PRF
ORGANISM: Pyrolobus fumarius
US-10-034-849-2

Query Match 100.0%; Score 4188; DB 12; Length 803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEVFTVLDSSYEYVGEKPOVIINGIAENGERVVLIDRSFRPYFALLAPGADPKQVQAQ 60
Db 1 MTEVFTVLDSSYEYVGEKPOVIINGIAENGERVVLIDRSFRPYFALLAPGADPKQVQAQ 60
Qy 61 RIRALSRPKSPIIGVEDDKKRYGRRRLRIRTVLPEAVREYRELKVNQDVEDVLEAD 120
Db 61 RIRALSRPKSPIIGVEDDKKRYGRRRLRIRTVLPEAVREYRELKVNQDVEDVLEAD 120
Qy 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYVLYKSRPEPLYGELAPTKLPDL 180
Db 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYVLYKSRPEPLYGELAPTKLPDL 180
Qy 181 RILAFDIEYVSKQSPRPREDPVIYIAVKTDDGDEVLFIAEGKDDKPIREFEYVKRYD 240
Db 181 RILAFDIEYVSKQSPRPREDPVIYIAVKTDDGDEVLFIAEGKDDKPIREFEYVKRYD 240
Qy 241 PDIIYGVNNHFMWPLRLRRARILGKLDVTRVGAEPITSVGHVSVGRANLVDLYXA 300
Db 241 PDIIYGVNNHFMWPLRLRRARILGKLDVTRVGAEPITSVGHVSVGRANLVDLYXA 300
Qy 301 EEMPEIKIKSLAEVAEYLCGMKKSERYIINMWEIPDYWDPKRRPLLQYARDVRAATYG 360
Db 301 EEMPEIKIKSLAEVAEYLCGMKKSERYIINMWEIPDYWDPKRRPLLQYARDVRAATYG 360
Qy 361 LAEKILPFAIQLSYVTGLPLDVGAMSVGFLEWYLIRAAFKMKEIVPNRVERPRETYRG 420
Db 361 LAEKILPFAIQLSYVTGLPLDVGAMSVGFLEWYLIRAAFKMKEIVPNRVERPRETYRG 420
Qy 421 AIVLEPLRGVHENIAVLDFSMTYPMIKYVNGPDLTVRGEKCGCGCMEAPEYKHF 480
Db 421 AIVLEPLRGVHENIAVLDFSMTYPMIKYVNGPDLTVRGEKCGCGCMEAPEYKHF 480
Qy 481 RCPGPFKTVLERLELRKRVRAEMKKYPPDSPEYRLDLDRQALVLANASVGMGSG 540
Db 481 RCPGPFKTVLERLELRKRVRAEMKKYPPDSPEYRLDLDRQALVLANASVGMGSG 540
Qy 541 ARWYCRECAKAVTAMGRHILRTAINIAKGLKXIVYDQDLSLVTVTPDKVENFKIKKE 600

Db 541 ARWYCRECAKAVTAMGRHILRTAINIAKGLKXIVYDQDLSLVTVTPDKVENFKIKKE 600
Qy 601 ELAFETIKLEKRYKRLFTTEAKRYAGLBDGRIDIVGEAVRGWGCCLAEVOTKVEIV 660
Db 601 ELAFETIKLEKRYKRLFTTEAKRYAGLBDGRIDIVGEAVRGWGCCLAEVOTKVEIV 660
Qy 661 LKTSVNRKAVYKRYKRYELBEKVPYIEKLVYKTLSTRLEETTEAPHYAAKRLSAG 720
Db 661 LKTSVNRKAVYKRYKRYELBEKVPYIEKLVYKTLSTRLEETTEAPHYAAKRLSAG 720
Qy 721 YRVSPGDKIGYIVYKGGGRISORAMPYFMVKDPSQIDVTYVVDHQQIIPALRILGYFGIT 780
Db 721 YRVSPGDKIGYIVYKGGGRISORAMPYFMVKDPSQIDVTYVVDHQQIIPALRILGYFGIT 780
Qy 781 EKKLKASATGOKTLFDFLAKSK 803
Db 781 EKKLKASATGOKTLFDFLAKSK 803

RESULT 6

US-10-034-621-2
Sequence 2, Application US/10034621
Patent No. US2002013297A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: CALLEN, Walter
APPLICANT: MATHER, Eric
TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY
TITLE OF INVENTION: AND METHODS OF USE
FILE REFERENCE: DIVER1350-2
CURRENT APPLICATION NUMBER: US/10/034, 621
CURRENT FILING DATE: 2000-12-21
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 803
TYPE: PRF
ORGANISM: Pyrolobus fumarius
US-10-034-621-2

Query Match 100.0%; Score 4188; DB 12; Length 803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEVFTVLDSSYEYVGEKPOVIINGIAENGERVVLIDRSFRPYFALLAPGADPKQVQAQ 60
Db 1 MTEVFTVLDSSYEYVGEKPOVIINGIAENGERVVLIDRSFRPYFALLAPGADPKQVQAQ 60
Qy 61 RIRALSRPKSPIIGVEDDKKRYGRRRLRIRTVLPEAVREYRELKVNQDVEDVLEAD 120
Db 61 RIRALSRPKSPIIGVEDDKKRYGRRRLRIRTVLPEAVREYRELKVNQDVEDVLEAD 120
Qy 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYVLYKSRPEPLYGELAPTKLPDL 180
Db 121 IRFAMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYVLYKSRPEPLYGELAPTKLPDL 180
Qy 181 RILAFDIEYVSKQSPRPREDPVIYIAVKTDDGDEVLFIAEGKDDKPIREFEYVKRYD 240
Db 181 RILAFDIEYVSKQSPRPREDPVIYIAVKTDDGDEVLFIAEGKDDKPIREFEYVKRYD 240
Qy 241 PDIIYGVNNHFMWPLRLRRARILGKLDVTRVGAEPITSVGHVSVGRANLVDLYXA 300
Db 241 PDIIYGVNNHFMWPLRLRRARILGKLDVTRVGAEPITSVGHVSVGRANLVDLYXA 300
Qy 301 EEMPEIKIKSLAEVAEYLCGMKKSERYIINMWEIPDYWDPKRRPLLQYARDVRAATYG 360
Db 301 EEMPEIKIKSLAEVAEYLCGMKKSERYIINMWEIPDYWDPKRRPLLQYARDVRAATYG 360
Qy 361 LAEKILPFAIQLSYVTGLPLDVGAMSVGFLEWYLIRAAFKMKEIVPNRVERPRETYRG 420

```

Db 361 LAEKILPFAIQLSVTGLPLDQVGAMSVGRLEWYLIRAAFKMKELVNNRVERPEETRG 420
Qy 421 AIVLEPLRGVHENTAVLDFSSMTYPMIMIKYNGPDTLVRGEGKCECCWEAPEVKHFR 480
Db 421 AIVLEPLRGVHENTAVLDFSSMTYPMIMIKYNGPDTLVRGEGKCECCWEAPEVKHFR 480
Qy 481 RCPGPFKTYLERLLELRKVRRAEMKKYPPDSPERYLLDEROKALKYLANASYGMGMSG 540
Db 481 RCPGPFKTYLERLLELRKVRRAEMKKYPPDSPERYLLDEROKALKYLANASYGMGMSG 540
Qy 541 ARMYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYDGTDSLFTYDPEKXENFIKIIE 600
Db 541 ARMYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYDGTDSLFTYDPEKXENFIKIIE 600
Qy 601 ELGFEIKLEKVKYKRLFFTEAKKRYAGLLEDGRIDIVGEFVARGMWCCLAKEVQIKVEIV 660
Db 601 ELGFEIKLEKVKYKRLFFTEAKKRYAGLLEDGRIDIVGEFVARGMWCCLAKEVQIKVEIV 660
Qy 661 LKTSBVNKAVEYVYRKIVKELEBEGKVPLEKLVIMKTLSEYTTAEAPHVVAAKRMLSAG 720
Db 661 LKTSBVNKAVEYVYRKIVKELEBEGKVPLEKLVIMKTLSEYTTAEAPHVVAAKRMLSAG 720
Qy 721 YRVSFGDKIGYVIVKGGGRISQRAWPFMYVDPQSQIDVTYYVDHQIIPALRIILGYFGIT 780
Db 721 YRVSFGDKIGYVIVKGGGRISQRAWPFMYVDPQSQIDVTYYVDHQIIPALRIILGYFGIT 780
Qy 781 EKKLKASATGQKTLFDFLAKSK 803
Db 781 EKKLKASATGQKTLFDFLAKSK 803

RESULT 7
US-09-948-369-16
; Sequence 16, Application US/09948369
; Patent No. US20020132243A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter
; APPLICANT: MATHEN, Eric
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS C
; FILE REFERENCE: DIVER1350-3
; CURRENT APPLICATION NUMBER: US/09/948,369
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US 09/656,309
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/391,340
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/907,166
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 803
; TYPE: PR
; ORGANISM: Pyrococcus furiosus
US-09-948-369-16

Query Match 99.4%; Score 4164; DB 10; Length 803;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 797; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MTEVFTVLDSYEVGKEPQVIITWGIANGERVVLLIDRSFRPYFYALLAPGADPKQVAQ 60
Db 1 MTEVFTVLDSYEVGKEPQVIITWGIANGERVVLLIDRSFRPYFYALLAPGADPKQVAQ 60
Qy 61 RIRALSRKSPITIGVEDDKKVFGRPRVLRIRTLPEAVNEYRLVKNVNGVEVLEAD 120
Db 61 RIRALSRKSPITIGVEDDKKVFGRPRVLRIRTLPEAVNEYRLVKNVNGVEVLEAD 120
Qy 121 IRFAMRYLIDHDLFPFTWYRVAEAPLENKMGFRVDKYLKSRPPLVGEALAPTKLPDL 180

```

```

Db 121 IRFAMRYLIDHDLFPFTWYRVAEAPLENKMGFRVDKYLKSRPPLVGEALAPTKLPDL 180
Qy 181 RIRALFIDIEVYSKQSPPEEDPVIIVIAVKTDDGDEVLFIAEGKDRKPIREPEVYKRYD 240
Db 181 RIRALFIDIEVYSKQSPPEEDPVIIVIAVKTDDGDEVLFIAEGKDRKPIREPEVYKRYD 240
Qy 241 PDIIIGVNNHFMPPYLIRBARILIGILKLDVTRRGAPPTSYVGHVSVPRLVNDLVDA 300
Db 241 PDIIIGVNNHFMPPYLIRBARILIGILKLDVTRRGAPPTSYVGHVSVPRLVNDLVDA 300
Qy 301 EEMPEIKISLEAEVAYELGVWKKSERVIIMWEIPDWDPPKPKRPLLOQYARDDVRATYG 360
Db 301 EEMPEIKISLEAEVAYELGVWKKSERVIIMWEIPDWDPPKPKRPLLOQYARDDVRATYG 360
Qy 361 LAEKILPFAIQLSVTGLPLDQVGAMSVGRLEWYLIRAAFKMKELVNNRVERPEETRG 420
Db 361 LAEKILPFAIQLSVTGLPLDQVGAMSVGRLEWYLIRAAFKMKELVNNRVERPEETRG 420
Qy 421 AIVLEPLRGVHENTAVLDFSSMTYPMIMIKYNGPDTLVRGEGKCECCWEAPEVKHFR 480
Db 421 AIVLEPLRGVHENTAVLDFSSMTYPMIMIKYNGPDTLVRGEGKCECCWEAPEVKHFR 480
Qy 481 RCPGPFKTYLERLLELRKVRRAEMKKYPPDSPERYLLDEROKALKYLANASYGMGMSG 540
Db 481 RCPGPFKTYLERLLELRKVRRAEMKKYPPDSPERYLLDEROKALKYLANASYGMGMSG 540
Qy 541 ARMYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYDGTDSLFTYDPEKXENFIKIIE 600
Db 541 ARMYCRECAKAVTAMGRHLIRTAIINARKLGLKVIYDGTDSLFTYDPEKXENFIKIIE 600
Qy 601 ELGFEIKLEKVKYKRLFFTEAKKRYAGLLEDGRIDIVGEFVARGMWCCLAKEVQIKVEIV 660
Db 601 ELGFEIKLEKVKYKRLFFTEAKKRYAGLLEDGRIDIVGEFVARGMWCCLAKEVQIKVEIV 660
Qy 661 LKTSBVNKAVEYVYRKIVKELEBEGKVPLEKLVIMKTLSEYTTAEAPHVVAAKRMLSAG 720
Db 661 LKTSBVNKAVEYVYRKIVKELEBEGKVPLEKLVIMKTLSEYTTAEAPHVVAAKRMLSAG 720
Qy 721 YRVSFGDKIGYVIVKGGGRISQRAWPFMYVDPQSQIDVTYYVDHQIIPALRIILGYFGIT 780
Db 721 YRVSFGDKIGYVIVKGGGRISQRAWPFMYVDPQSQIDVTYYVDHQIIPALRIILGYFGIT 780
Qy 781 EKKLKASATGQKTLFDFLAKSK 803
Db 781 EKKLKASATGQKTLFDFLAKSK 803

RESULT 8
US-09-391-340-6
; Sequence 6, Application US/09391340A
; Patent No. US20020013455A1
; GENERAL INFORMATION:
; APPLICANT: Calleen, Walter
; APPLICANT: Mathen, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/09/391,340A
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/907,166
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 788
; TYPE: PR
; ORGANISM: Archaeoglobus fulgidus
US-09-391-340-6

Query Match 37.3%; Score 1563; DB 10; Length 788;
Best Local Similarity 41.3%; Pred. No. 5; Indels 54; Gaps 18;
Matches 338; Conservative 154; Mismatches 273; Indels 54; Gaps 18;
Qy 1 MTEVFTVLDSYEVGKEPQVIITWGIANGERVVLLIDRSFRPYFYALLAPGADPKQVAQ 60

```

Wed Apr 23 08:09:12 2003

us-10-034-849-2.rapb

Page 6

```
Db 1 MIKKGWLLDADYITENDRAVIRLWCKDEEG-IFIAVDSFQPYFALKEEGITAEIVK 59
Qy 61 -RIRALSRPKSPIIGVEDDKRYFGRPRVLRIRTVLPEAVREYRELVKNVGVEDVLEA 119
Db 60 IKVQTKKEVITP-LKVEETAKNLGREVVEFKIYARHPQHVPLKREVSQ---YLEIREA 115
Qy 120 DIRFAMRYLIDHDLFPFTWYRVEAEPLNKGSPRDVYLVK--SRPEPLYGALAPYKL 177
Db 116 DIPFAYRYLIDKNLACMDGVILEGVERREK-GLRCYEIKRIERDSQD-----F 163
Qy 178 PDRLIAPDIEVYKSGSPRPREDPVYIAVKTDDGDEVLFIAGKODRKPIREVEYVK 237
Db 164 PELKWAHPDCMLSEVGMDEPEKDPPIIVISISGVEEIL---NGDNRELLTRFKIIR 220
Qy 238 RYDDPIIVGYNHNPFPYLLRRARILGILQDVTRVGAEPSTSVGHVSPGRNLVDLY 297
Db 221 DIDPDIIVGYNQDSFDPWPIKRAEKLRYKLDIGR--DRSELAI RGRPKIAGRLNVDLY 278
Qy 298 DYAEEMPEIKIKSLAEVAEYLVGMKSSRVLIINWMEIPDYWDDPKRPLLQYARDVRA 357
Db 279 DIAMRSIDVYKXKLENVAEFLG--KKEIADIEAKDIYKHWTS--GDRESVIKYSRODILH 335
Qy 358 TYGLAEKILPFAIOLSVYTGPLDPOVAMSVGRLEWYLIRAAFKMKELVPRNVERPEET 417
Db 336 TYPIAEELLPMHYELSRMIRIPLDDVTRSGRGKQVEWLLSEAHKUGELAPNPREM-ADS 394
Qy 418 YRGAIVLEPLRGVHNIIVLDFSSMYPNIMIKYVGPDTLVPRPGEKSGEGCWEAPEYKH 477
Db 395 YEGAFVLEPARGLHENVICLDFASMPISIMISPTLVI--GKDDCNV--APEVGH 450
Qy 478 RFRCPGPFKTVLERLELRKVRAMKXPPSPPEYRLDERQKALVLANASGYWG 537
Db 451 KFRGHPGPFKRIKLIKREBEIKKMKITDYNSEYKLDIKQATLVLTNSFYGTG 510
Qy 538 WSGARWYCECAKAVTANGRHILRTAINIARKLGKVIYGDTSLFVTVYDPEKVENFIKI 597
Db 511 WSLARWYCECAKAVTANGRHILRTAINIARKLGKVIYGDTSLFVTVYDPEKVENFIKI 597
Qy 598 IK-----BELGFEIKLEKRYKRLFTFAKKRYAGLLEDGRIDIVGFPAVNGDCELAK 650
Db 571 VKKLIGLSEEMPIQIEIDEYEETIFPVE-KRKYAGLTQDGRITIVKLEVRGDMCELAK 629
Qy 651 EVQTKVEIVLKTSEVNKAVEYRKIVKELEBGRVPLEKVIWKTLSKLEEYTTAPHV 710
Db 630 KIQGVIEIILKEKNPEKAAYVGVIEIEIKAKIPLIEDYIYKGLTRKPSKYESMOAHV 689
Qy 711 VAAKRLMSAGYRVPDGIYVIVKGGRIISORAMPYFMYKD-----PSQIDV 758
Db 690 KAAMKAARIGIYVITIGSKYGVTVTKGVNIGDRAFPDLIEDPDGEVITDLDGNKYKIDK 749
Qy 759 TYVVDHIIIPALRIIGVYGTIEKKLKASATGOKTLFDF 797
Db 750 EYIDHQVLPSTVRLIRERFGYTEAQKGAAB-QQTLDAF 787

RESULT 9
US-09-948-369-6
; Sequence 6, Application US/09948369
; Patent No. US20020132243A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter
; APPLICANT: MAYHUR, Eric
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS C
; TITLE OF INVENTION: THERMOF
; FILE REFERENCE: DIVER1350-3
; CURRENT APPLICATION NUMBER: US/09/948,369
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US 09/656,309
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/391,340
; PRIOR FILING DATE: 1999-09-07
```

```
PRIOR APPLICATION NUMBER: US 08/907,166
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 788
; TYPE: PRT
; ORGANISM: Archaeoglobus lithotrophicus
US-09-948-369-6

Query Match 37.3%; Score 1563; DB 10; Length 788;
Bseq Local Similarity 41.3%; Freq. No. 5.7e-120; Indels 54; Gaps 18;
Matches 338; Conservative 154; Mismatches 273;

Qy 1 MTEVFTVLDSSYEYVKEPQVITWGIANGERRVLLIDRSFRPYFVALLAPGADPKQVAQ 60
Db 1 MIKKGWLLDADYITENDRAVIRLWCKDEEG-IFIAVDSFQPYFALKEEGITAEIVK 59
Qy 61 -RIRALSRPKSPIIGVEDDKRYFGRPRVLRIRTVLPEAVREYRELVKNVGVEDVLEA 119
Db 60 IKVQTKKEVITP-LKVEETAKNLGREVVEFKIYARHPQHVPLKREVSQ---YLEIREA 115
Qy 120 DIRFAMRYLIDHDLFPFTWYRVEAEPLNKGSPRDVYLVK--SRPEPLYGALAPYKL 177
Db 116 DIPFAYRYLIDKNLACMDGVILEGVERREK-GLRCYEIKRIERDSQD-----F 163
Qy 178 PDRLIAPDIEVYKSGSPRPREDPVYIAVKTDDGDEVLFIAGKODRKPIREVEYVK 237
Db 164 PELKWAHPDCMLSEVGMDEPEKDPPIIVISISGVEEIL---NGDNRELLTRFKIIR 220
Qy 238 RYDDPIIVGYNHNPFPYLLRRARILGILQDVTRVGAEPSTSVGHVSPGRNLVDLY 297
Db 221 DIDPDIIVGYNQDSFDPWPIKRAEKLRYKLDIGR--DRSELAI RGRPKIAGRLNVDLY 278
Qy 298 DYAEEMPEIKIKSLAEVAEYLVGMKSSRVLIINWMEIPDYWDDPKRPLLQYARDVRA 357
Db 279 DIAMRSIDVYKXKLENVAEFLG--KKEIADIEAKDIYKHWTS--GDRESVIKYSRODILH 335
Qy 358 TYGLAEKILPFAIOLSVYTGPLDPOVAMSVGRLEWYLIRAAFKMKELVPRNVERPEET 417
Db 336 TYPIAEELLPMHYELSRMIRIPLDDVTRSGRGKQVEWLLSEAHKUGELAPNPREM-ADS 394
Qy 418 YRGAIVLEPLRGVHNIIVLDFSSMYPNIMIKYVGPDTLVPRPGEKSGEGCWEAPEYKH 477
Db 395 YEGAFVLEPARGLHENVICLDFASMPISIMISPTLVI--GKDDCNV--APEVGH 450
Qy 478 RFRCPGPFKTVLERLELRKVRAMKXPPSPPEYRLDERQKALVLANASGYWG 537
Db 451 KFRGHPGPFKRIKLIKREBEIKKMKITDYNSEYKLDIKQATLVLTNSFYGTG 510
Qy 538 WSGARWYCECAKAVTANGRHILRTAINIARKLGKVIYGDTSLFVTVYDPEKVENFIKI 597
Db 511 WSLARWYCECAKAVTANGRHILRTAINIARKLGKVIYGDTSLFVTVYDPEKVENFIKI 597
Qy 598 IK-----BELGFEIKLEKRYKRLFTFAKKRYAGLLEDGRIDIVGFPAVNGDCELAK 650
Db 571 VKKLIGLSEEMPIQIEIDEYEETIFPVE-KRKYAGLTQDGRITIVKLEVRGDMCELAK 629
Qy 651 EVQTKVEIVLKTSEVNKAVEYRKIVKELEBGRVPLEKVIWKTLSKLEEYTTAPHV 710
Db 630 KIQGVIEIILKEKNPEKAAYVGVIEIEIKAKIPLIEDYIYKGLTRKPSKYESMOAHV 689
Qy 711 VAAKRLMSAGYRVPDGIYVIVKGGRIISORAMPYFMYKD-----PSQIDV 758
Db 690 KAAMKAARIGIYVITIGSKYGVTVTKGVNIGDRAFPDLIEDPDGEVITDLDGNKYKIDK 749
Qy 759 TYVVDHIIIPALRIIGVYGTIEKKLKASATGOKTLFDF 797
Db 750 EYIDHQVLPSTVRLIRERFGYTEAQKGAAB-QQTLDAF 787

RESULT 10
US-09-852-922-2
```


Sequence 2, Application US/09862922
Patent No. US20020076768A1
GENERAL INFORMATION:
APPLICANT: TOYO BOSEKI KABUSHIKI KAISHA
TITLE OF INVENTION: MODIFIED THERMOSTABLE DNA POLYMERASE
FILE REFERENCE: 000053
CURRENT APPLICATION NUMBER: US/09/852,922
CURRENT FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 2000-138796
PRIOR FILING DATE: 2000-05-11
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 774
TYPE: PRT
ORGANISM: *Pyrococcus kodakarensis*
US-09-852-922-2

Query Match	29.7%	Score 1242.5	DB 10	Length 774
Best Local Similarity	36.7%	Pred. No. 1.3e-93		
Matches 304	Conservative 144	Mismatches 288	Indels 93	Gaps 22

QY	8	VLDSSYEWGVEPOVYIWMGIMENGERVVLIDRSRPYFVALLAAGAPKOV----	AQRIR	63
Db	2	ILUDDYITDEOKPVIARIFK-KENGEFKIETDRTPEPFIYALKDOSAIEEKKYITABRHG		60
QY	64	ALSPPKSLIGVEDDKRKYFPRPRVIRIRTVLPEAVREYELVKANDGVEDVLEADIRF		123
Db	61	TVTVTVKR----VEKQKKFKLGRPVEWMKLYPTHQODVPALRDKIRHPAVIDIYEYDIPF		116
QY	124	AMRVLIDHDLFPFTWYRVEAPELENGKGFVVDKYLYKSRPEPYLGEALAPTCLPDLRIL		183
Db	117	AKRVLIDGLV-----PMEGD-----EELKML		138
QY	184	AFDIYEVSKQSPRPERDPVIVIAVKTDDGEVL-----FLAEKGDKRPIREVEY		235
Db	139	AFDIETLYHBEEBEAF-OPIMISYADEBKAVITMKVNDLPYDVVSTEBEMIKRFLRV		197
QY	236	VKRYDPLIIYGVYNNHEDPWLRLRRARILGIKLVTRRVGAEPPTSVMH-----VSVGRL		292
Db	198	VKEKDOPVLIIYVNDPNDFALYKRRCEKLGINPALGND-GSEPKIQMGDRPAVEVAGRI		256
QY	293	NVDLYDAEEMPEIKISLEBAVLYGMKKSERYIINMEIPIYWDPKKRPILLOYAR		352
Db	257	HFDIYPVIRRTINLPYTLAEVYEAIVEGQPK-EKYVAE--EITTAMEGTEMLERVAYSM		313
QY	353	DDVATYGLAKKILPRAIQUSVYVGLPUDQVGMASVGERLEWYLIRAAFKKELVPRVVE		412
Db	314	EDAVYTYELGKEPFLPMEHQSRLLGQSIQMDVSRSSGTGLVEMFLLRKAYERNELAPKPKD		373
QY	413	RPE-----ETRGAILVLEPLRGVENIAVLDFSSMYPNIMKYVGVGDTLVPRPEKGC		467
Db	374	EKELARRQSVREGGVYVEPERGMLMENIYVLDPRSLYSIIITHNVSDTLNRE-----		426
QY	468	GCWE--APEVKRPRRCRCPGFPTVLERLLERKRVAAEMKKYPDPSPEYRLDEROKA		524
Db	427	GCKEYDVAPOVGHRCFDKPFIFISLLGDLLEEROKIKKKWKA-TIDPIERKLLDYQRA		489
QY	525	LKVLNAYSXGMGMSGARWYCREGAKAVTANGRLITAI-NIARUKGLKVIYDSDTSLF		583
Db	486	IKILANSTYGIYGARAPAWYCKECAESVYTAANGREYITMTJKEIBEKGFVYISDITGFF		545
QY	584	VTV---DPEKVE---NFIKIIKEEL--GFPIKLEKYKRLFTEAKKRVAGLLEDGRID		634
Db	546	ATIPGADAEIYKXKKAMELKITYNKLPRALELEIEGPIKGFY--TKKRIAVIDEBSKIT		604
QY	635	IIVGEPAVAGDMCELAKEVQTVAVEIVLKTSFVNKAVEVVRKIYKELEBGRVPIEKLVIWK		694
Db	605	TRGEIIVARDSLSIAKETQARVLBALLKDGQVEKAVRIYKEVTEKLSKYVEPPEKLVTHE		664
QY	695	TLISRLBEYTTAEHHVNAKMLBAGRVASGDKIGVIVYVGGRIISQRAPIYMWVQDS		754
Db	665	QITDULQOYKATGPBVAVAKRLAGVIRGTVISYIVLVGSGRISIDRPIR-FDEDDPT		723

```
QY 755 --QIDVTYYVDHQIIPALRILGFGITTEKKLKASATGQKTLFDPLAKK 801
  : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 724 KHKYDAEYIENQVLPAVERILRAFGRKEDLRYQKTRQVGLSAWLKPK 772
```

RESULT 11

US-09-803-165-34
; Sequence 34, Application US/09803165

```

! TYPE: PRT
! ORGANISM: T. aggregans
US-09-803-165-34

```

Query Match	28.5%	Score 1194	DB 10	Length 774
Best Local Similarity	35.1%	Pred. No. 1.2e-89		
Matches 291	Conservative 148	Mismatches 298	Indels 92	Gaps
QY	8	VLSVSVEVGKEPOVILIMJGIAENBERVVLIDRSFRPFYALLAPGADPKOAGIRALSR	67	
Db	2	IFDIDYITTKOGKPIIRIFK-KENEFRIEIDPHQPIITLYLLDSDS---AIDEIKAIKG	56	
QY	68	PK-SPIIGVEDD--KRTYGRPRRVLRITVLEPAVREYRELKONVGVEDVLEADIRF	1223	
Db	57	ERHGKIVRVVDVAVKKKKFFGRDVEVWKLLFEHPDVPALRGKIREHPAVIDIYEYDIPF	1161	
QY	124	AKRLIHDHLEPFTWYRVVEAPLENNKGFRVDKYLKRSRPEPLVGEALAPTKLPDLRIL	1893	
Db	117	AKRLIDKGLI-----PWEGD-----EELKIM	1383	
QY	184	AFDIEVYSKQSGPRPERDPVIVIAVKTDDGDEVL-----FLAEGKDRKPIREFVEY	2359	
Db	139	AFDIEYTFYHEDBERGKGE-IIMISYADEEBEARVITWKINIDLPYDVVDSNREMIKRFVQI	1971	
QY	236	VKRVDPPIIYGNNHHNDWVYLRLBRALIGIKLDVTRVGAEPRTSHG-----HVSVPG	2902	
Db	198	VREKDPVLLITNGDNDLPYLIRGAKTLGVTL-LGRDKYHEBPPKIRHMGDSFAVEIKG	2561	
QY	291	RLNVDLYVAEEMBEIKISLSEVYAEYLGVWKSERYIINWBEIPDYWDDPKKRPLLQY	3560	
Db	257	RIHFDELPPVVRRTITLPTYTLEAYEAVLGKTSK---LGAEBTALIMETESWKKLAQY	3113	
QY	351	ARDVNRVTYGAELILFPALQLSVYVGLPDDQVAGMVGSRLEBYVLLRAAFKMKELVPR	4110	
Db	314	SMEARATYELGKFFPEMEABLAQLDQSWVDVSRSTGMVLEVYTLIRVAYERNELAPNK	3737	
QY	411	VERDE-----ETVGAIVLEPLRGHENIAVLDFSSMYENIMKYVNGDTLVRPGEK	4647	
Db	374	PDEEYVRRRLRTTYLGGVYKBERGMENTITVDFRCLPSIITVHNVSOTLIEREG---	4303	
QY	465	GEQGCWE-ABEVKGRFRRCRPGFKVYLELLBELRKVVRABMKYPPDSSEYRLDEROK	5223	
Db	431	--CANYVAPVGVKFFCKDPGGLPSPISLTGELITRQEIKKKKMA-TIDPLEKMLDYROR	4877	
QY	524	ALKYLANASGYMGWSGARWYCECAAVAVAMGRHLIRTAI-NIARKTLGLKVIYGDYDSL	5823	

Db 488 AVKHAASYGWMGYRKARWYSEKCAESVTAMGRHYIEMTIKEEKGFKVLVADTGC 547
 QY 563 FVTYDPRKVENFKIIEKEIGF-----EIKLEKVVYKQLFPTAKKRYVGLLEDGRI 633
 Db 548 YATIPGKEPTIKKAKKEFKYIKNSKLGLELEBEGYIRGFV- AKRYVAVIDEERI 606
 QY 634 DIVFEAVRGDMCELAKEVQTRVVEIUKTSEVNAVEYRKIVKELEBEGYIRKIVYI 693
 Db 607 TTRLEAVRRDMSKIAKEVQAVLEAIKEDSVKAVEIVKDVVEIACYVPELEKVIH 666
 QY 694 KTLSEKLEEYTTAPPHVAAKMLSAGYVSPGDKIGYIVKGGRIISORAMPYVNDP 753
 Db 667 EOIIRDLSEYKAIQPHVAIAKRLAKGIRPGTIIISYIIRSGSKISDRV- ILLSEYDP 725
 QY 754 S-QIDVTYVVDHIIIPALRIILGYGITEKKLAKSATGOKTLPEDLAK 800
 Db 726 KKHKXDPDYIENQVLPVILRIIEAFGRKEDLKYGSSKQVGDAMLKK 774

RESULT 12

Sequence 30, Application US/10027806
 Patent No. US20020160476A1
 GENERAL INFORMATION:
 APPLICANT: Swanson, Ronald V.
 APPLICANT: Feldman, Robert A.
 APPLICANT: Schleper, Christa
 TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 FILE REFERENCE: DCOIP.002A
 CURRENT APPLICATION NUMBER: US/10/027,806
 CURRENT FILING DATE: 2001-12-21
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
 PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
 NUMBER OF SEQ ID NOS: 123
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 30
 LENGTH: 844
 TYPE: PRT
 ORGANISM: Cenarchaeum symbiosum
 US-10-027-806-30

Query Match 13.5%; Score 564.5; DB 9; Length 844;
 Best Local Similarity 25.9%; Pred. No. 8.3e-38;
 Matches 192; Conservative 127; Mismatches 300; Indels 122; Gaps 25;

Db 108 KNVGVEDVLEADIRFAMRYLIIDHDLFPTWYRVEA-----EPLEKMGFRVDKVVYL 160
 Db 106 KSIKIMDTWESDIKYENYLYDKSLVGRYYSVSGKVIPLHMPISDEV-----KLMALK 160
 Db 161 KSRPELYGEALAPTK-----LPDLRIIAFDIEVYSKQG--SPRPERD 201
 Db 161 SLIMDKVDEGMADRKEREPRFIAGMADLNPRIPIRIRLSFDIEVSEBGRIPDKISDR 220
 QY 202 PVIYIAVKTDDGDEVLFT-----AEKDKRKPIREFVEYVYKRYDP-----DIIV 245
 Db 221 RVTAVGPAATDGLKQVFLRSGAESEBENVTGVEVYVYDEADIMIDALSVIGSYPFL 280
 QY 246 GYNNHFWPMLYLRARILIGIK-LDVTIRVGAEPSTVHG-HVSV---GRLVNDLYDVA 300
 Db 281 TYNGDDFDMPTMLNRRARLGVSDSDIPLYMWRDATALRHGVHLDLRTFSNRSFOLYFA 340
 QY 301 EEMPEIKIKSLEVAEYLVGMKSERVIINMWEIPDYWDPKKRPILL-----QYARDVR 356
 Db 341 AKYTDYSINSVTKA--MLGEK-----VDY--GVKGLDITLYQTANYCYHDAR 384
 QY 357 ATYGLA-----EKLIPAIOLSVYVGLPLDOYGANSVGRLEWYLIAPAKMEIAPNVE 412
 Db 385 LTLELSTFGNEIIMDLVVTISRIAMPIDMSRMGVSQWIRSLYIEHQRALIPRDE 444
 QY 413 R-----PEETRGAIYVEPLRGVHENIAVLDSMSSYRPIIMIKYVAGDTLYR 459
 Db 445 LEGSREVSNDAYIKDKKFRGLVVEPEEGIHFPDYVMDPASYLSIIRKYNLSYEV-- 502

QY 460 PGEKC--GEGCWEADEVKRRFRCPPEFKTVLERLLELRKVRABMKYPPDSFEYRL 517
 Db 503 ---RCHVACCKKNITPDINHWVCTKNNGLTSMIISLDELVNYYKSLSTSITSE--- 556
 QY 518 LDERQ-----KALKVLNANSYGMGSGARWYCEKAKAVTANGRLHIRTAINIRKUG 571
 Db 557 --ORQOYTVISQALKVLNANSYGMGAEIIFPLYPAAEATTAAGRYIINQTIISHECQG 614
 QY 572 LKVIYDSTLSLEVYVYDE--KVENFKIIEKEIGFEIKLEKVVYKQLFPTAKKRYVAGLE 629
 Db 615 VRVLYGDDTSLEPIK-DEPERQIHEIYEHAKKEHGEVLEVDREYRVVLSNRKNYFGYTR 673
 QY 630 DGRIDIVGEAVRGDMCELAKEVQTRVVEI--VLTSEVNAKAVEYRKIV---KELE 682
 Db 674 AGKVDPKGLTGKKSHTPEPIKELFYSILDILSGVESEDEPESAKMRSKAIACGRLE 733
 QY 683 GKVPLEKVIWKTLSKLEEYTTAPPHVAAKRMISAGYVSPGDKIGYIVKGGRIISQ 742
 Db 734 RQIPLVNDLAFNMISKAPSEYVKTVPQHIRARLLENAREVYKGDIIISYKVNANKGVK- 792
 QY 743 RAMPYRMKDSQIDVTYYVD 763
 Db 793 ---FVENAR-AGEVDTSKYLE 809

RESULT 13

US-10-034-623-30
 Sequence 30, Application US/10034623
 Publication No. US20020198365A1
 GENERAL INFORMATION:
 APPLICANT: Swanson, Ronald V.
 APPLICANT: Feldman, Robert A.
 APPLICANT: Schleper, Christa
 TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 FILE REFERENCE: DCOIP.002A
 CURRENT APPLICATION NUMBER: US/10/034,623
 CURRENT FILING DATE: 2001-12-21
 PRIOR APPLICATION NUMBER: 09/408,020
 PRIOR FILING DATE: 1999-09-29
 PRIOR APPLICATION NUMBER: 60/102,294
 PRIOR FILING DATE: 1998-09-29
 NUMBER OF SEQ ID NOS: 123
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 30
 LENGTH: 844
 TYPE: PRT
 ORGANISM: Cenarchaeum symbiosum
 US-10-034-623-30

Query Match 13.5%; Score 564.5; DB 9; Length 844;
 Best Local Similarity 25.9%; Pred. No. 8.3e-38;
 Matches 192; Conservative 127; Mismatches 300; Indels 122; Gaps 25;

QY 108 KNVGVEDVLEADIRFAMRYLIIDHDLFPTWYRVEA-----EPLEKMGFRVDKVVYL 160
 Db 106 KSIKIMDTWESDIKYENYLYDKSLVGRYYSVSGKVIPLHMPISDEV-----KLMALK 160
 Db 161 KSRPELYGEALAPTK-----LPDLRIIAFDIEVYSKQG--SPRPERD 201
 Db 161 SLIMDKVDEGMADRKEREPRFIAGMADLNPRIPIRIRLSFDIEVSEBGRIPDKISDR 220
 QY 202 PVIYIAVKTDDGDEVLFT-----AEKDKRKPIREFVEYVYKRYDP-----DIIV 245
 Db 221 RVTAVGPAATDGLKQVFLRSGAESEBENVTGVEVYVYDEADIMIDALSVIGSYPFL 280
 QY 246 GYNNHFWPMLYLRARILIGIK-LDVTIRVGAEPSTVHG-HVSV---GRLVNDLYDVA 300
 Db 281 TYNGDDFDMPTMLNRRARLGVSDSDIPLYMWRDATALRHGVHLDLRTFSNRSFOLYFA 340
 QY 301 EEMPEIKIKSLEVAEYLVGMKSERVIINMWEIPDYWDPKKRPILL-----QYARDVR 356
 Db 341 AKYTDYSINSVTKA--MLGEK-----VDY--GVKGLDITLYQTANYCYHDAR 384

QY 357 ATYGLA-----EKILPFAIQLSVYTGCLPLDOVGAMSVGFRLBWLIRAAFKMKELVPRVE 412
 Db 385 LTLELSTFGNEILMDLLVTSRIARMPIDDMSRMGVQWISLSLYEHRQNNALIPRDE 444
 QY 413 R-----PEETRGALVLEPLRGVHENIAVLDPSSMYPNIMIKYNGPDTLVR 459
 Db 445 LEGRSREVNDVAIKDKKFRGGLVVEPEGHIFDVTAMDFASTLPSIIRKRNLSYETV-- 502
 QY 460 PGEKC--GEGCCWAPBVKRFRCPGPFKTVLERLELRKRVAAEMKKYPPDPSPEYRL 517
 Db 503 ---RCVHAECCKNTIPDTHNVCTKNNGLTSMIIGSLDLRNVYKSLKSTSTIEE--- 556
 QY 518 LDERQ-----KALKVLANASGYGWSGARWYCRECAKAVTAMGRHLIRTAIINARKLG 571
 Db 557 --QROQYTVISQALKVLANASGYGWSGARWYCRECAKAVTAMGRHLIRTAIINARKLG 614
 QY 572 LKVIYGDTSLFVTVYDPE--KVENFIKIIKEBELFEIKLEKVKYKRLPFTAKKRYAGLLE 629
 Db 615 VRVLYGDTDSLFIK--DPEBRQIHEIVEHAKKEHVELEVDKEXRYVVLNKRKNFVTVR 673
 QY 630 DGRIDIVGFEAVRGDMWCLAKEVOTKVEI---VLKTSVNKAVEYRKIV---KELEE 682
 Db 674 AGKDVVKGLTGKKSHTPPFIKELFYSLLDILSGVSEDEPESAKRISKAIACGKRLKE 733
 QY 683 GKVDIEKLVIMKTLKRLSEYTTAPHVAAKRLMAGYVSPGDKIGYIVYKGGGRISQ 742
 Db 734 RQILVLDLAFVNMISKAPSEYVKTVPQHIRARLLENAREVKGDIIISYVNMKTVGK- 792
 QY 743 RAMEFYVWKDPSQIDVTYVVD 763
 Db 793 ---EVEAR--AGEVDTSKYLE 809

RESULT 14
 US-10-027-801-30
 ; Sequence 30, Application US/10027801
 ; Publication No. US20030054364A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Swanson, Ronald V.
 ; APPLICANT: Feldman, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 ; FILE REFERENCE: DCOIP.002A
 ; CURRENT APPLICATION NUMBER: US/10/027,801
 ; PRIOR FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 30
 ; LENGTH: 844
 ; TYPE: PRF
 ; ORGANISM: Cenarchaeum symbiosum
 US-10-027-801-30

Query Match 13.5%; Score 564.5; DB 9; Length 844;
 Best Local Similarity 25.9%; Pred. No. 8.3e-38;
 Matches 192; Conservative 127; Mismatches 300; Indels 122; Gaps 25;
 QY 108 KNVDSGEVDLEADIRFAMRYLIDHDLFPFTWYRVEA-----EPLNMGGRVVDKYLIV 160
 Db 106 KSRINIMTWESDIKYENYLYDKSLVGRYYSVSGKVIPIHDMISDBV-----KMLK 160
 QY 161 KSRPEPLGALAPTK-----LPLRLIAPDIEVYSKQG--SPRPERD 201
 Db 161 SLMDKVVDEGAKKERREFFIAGWADLLNCPPIRIRLSFQIEVDSEGRIPDKISDR 220
 QY 202 FYIVAVATDDGDVLFPI-----AEGKDRKPIREFVYVYKYP-----DIIV 245
 Db 221 RVTAVGPAATDGLKQVFLVRSAGEGEGVTVGVEVVFDEKADIRDALSVISGYPFVL 280
 QY 246 GYNNHFMPPYLRLRARIIGIK-LDVTFRVGAEPPTTSVHG--HVSVP---GRILVLDLYDYA 300

Db 281 TYNGDDFDMPYMLNRRARLGVSDSDIPLVMMRDSATLRHGHLIDYRTFSNRSPQLYFA 340
 QY 301 EEMPEIKISLEAEVGLVWKKSERYIIMWELPDYWDPKKRPLL---QYARDVR 356
 Db 341 AKYDYSLNSVTXA--MIGGK-----VDY--GVKLDLTLYOTANYCYHDAR 384
 QY 357 ATYGLA-----EKILPFAIQLSVYTGCLPLDOVGAMSVGFRLBWLIRAAFKMKELVPRVE 412
 Db 385 LTLELSTFGNEILMDLLVTSRIARMPIDDMSRMGVQWISLSLYEHRQNNALIPRDE 444
 QY 413 R-----PEETRGALVLEPLRGVHENIAVLDPSSMYPNIMIKYNGPDTLVR 459
 Db 445 LEGRSREVNDVAIKDKKFRGGLVVEPEGHIFDVTAMDFASTLPSIIRKRNLSYETV-- 502
 QY 460 PGEKC--GEGCCWAPBVKRFRCPGPFKTVLERLELRKRVAAEMKKYPPDPSPEYRL 517
 Db 503 ---RCVHAECCKNTIPDTHNVCTKNNGLTSMIIGSLDLRNVYKSLKSTSTIEE--- 556
 QY 518 LDERQ-----KALKVLANASGYGWSGARWYCRECAKAVTAMGRHLIRTAIINARKLG 571
 Db 557 --QROQYTVISQALKVLANASGYGWSGARWYCRECAKAVTAMGRHLIRTAIINARKLG 614
 QY 572 LKVIYGDTSLFVTVYDPE--KVENFIKIIKEBELFEIKLEKVKYKRLPFTAKKRYAGLLE 629
 Db 615 VRVLYGDTDSLFIK--DPEBRQIHEIVEHAKKEHVELEVDKEXRYVVLNKRKNFVTVR 673
 QY 630 DGRIDIVGFEAVRGDMWCLAKEVOTKVEI---VLKTSVNKAVEYRKIV---KELEE 682
 Db 674 AGKDVVKGLTGKKSHTPPFIKELFYSLLDILSGVSEDEPESAKRISKAIACGKRLKE 733
 QY 683 GKVDIEKLVIMKTLKRLSEYTTAPHVAAKRLMAGYVSPGDKIGYIVYKGGGRISQ 742
 Db 734 RQILVLDLAFVNMISKAPSEYVKTVPQHIRARLLENAREVKGDIIISYVNMKTVGK- 792
 QY 743 RAMEFYVWKDPSQIDVTYVVD 763
 Db 793 ---EVEAR--AGEVDTSKYLE 809

RESULT 15
 US-10-027-806-62
 ; Sequence 62, Application US/10027806
 ; Patent No. US20020160476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Swanson, Ronald V.
 ; APPLICANT: Feldman, Robert A.
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 ; FILE REFERENCE: DCOIP.002A
 ; CURRENT APPLICATION NUMBER: US/10/027,806
 ; PRIOR FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 62
 ; LENGTH: 845
 ; TYPE: PRF
 ; ORGANISM: Cenarchaeum symbiosum
 US-10-027-806-62

Query Match 13.2%; Score 552; DB 9; Length 845;
 Best Local Similarity 26.0%; Pred. No. 8.9e-37;
 Matches 209; Conservative 133; Mismatches 339; Indels 124; Gaps 27;
 QY 42 RPYVYALLAPGADPKQVAGRIARLSPKSPIGVEDDKRYKFGRR--VLRIRTVLPEA 99
 Db 47 KPYCYTRLPSR-----LGFSGREVDLGIQVNRHDLIADKVPVSKTIVSDPLA 97
 QY 100 VREYRELIVKNVGEVDLEADIRFAMRYLIDHDLFPFTWYRV--EAPLENMGGRVVDK 156
 Db 98 IGTGHS--EKSIRINVIDWESDIKYENYLYDAGLVGRYYSVSGEVIPIHDMISDBV-K 155

Wed Apr 23 08:09:12 2003

us-10-034-849-2.rapb

Page 10

QY 157 VYLVKSRPEPLYGALAPTK-----LPDDRILAFDIEVYSKOG-----SP 196
DB 156 IALKSLIMDLKIDEGWADRKEPERFIAGMADLNLQPIPRIRRSFDIEVDSSEGRIPDAK 215
QY 197 RPERDPYIVIAVKTDDGDEVLFA---EGKDDRKPIREFEVEYKRPDI-----244
DB 216 ISDRRYAVAGFAITDGLRVLVUKSGADEGANBVTBGEVVFYDEKENDMIRDALIIG 275
QY 245 -----VGYNNHEDWYLLRRARITGI-KLDVTRVGAEPITSG-HVSVP--GRINV 294
DB 276 SYPFVLTINGGDDPDMYNNRRKRLGADSDIPLYMMDSATLRHGVHLDLRTPSNRSF 335
QY 295 DLYDYAEEMPEIKIKSLIEVAVYLGWKKSERVIINMWEIPDYWDDPKRPILLQYARD 354
DB 336 QLYAFAAKYTDYSLNSVSKA--MDEGKYDYVSLG--DLTY-----OTANYCYHD 383
QY 355 VRATYGLA---EKILPFAIOLSYVTGLPLDQVGAMSVGFLEMYLIRAAFOMKELVNR 410
DB 384 ARLTLELSTFGNEILMDLIVTSRIARMPIDMSRMGVSQMIRSLIYEHQRNALIPRR 443
DB 411 VER-----PEETYGAIVLEPLGVEHENIAVADPSSMYPNIMIKYNGPDTL 457
DB 444 DELEKRSQVSNDAVVKDKKFRGGLVVEPEGIHFDVTVMDFASLYPSIIKYRNLSEYEV 503
QY 458 --VAPGEKCGECGCEAPFVKRFRRCPPGFVTLERLLELRKRVRAEMKKYPPDSPEY 515
DB 504 RCVAF-----ECKRNTIPIPTNHVCTKNNGLTSMIIGSLDLKAVNYKSLSKQSITHE 557
QY 516 RLUDERQ-----KALKVLANASYGMGMSGARWYCHRECAKATAMGRHLITAINIARK 569
DB 558 ----QROQYVISQALKVLANASYGMGAEIIFPIYLPAAEATTAVGRITIMQISHCQ 613
QY 570 LGLKVIYGDPSLFTYDPE--XVENFIKIKEELGPEIKLEKYRQLFTEAKRYAGL 627
DB 614 MGVAVLYGDDLSLPIK-NPEERQIHDIYEHAKKEHGELEVDKEYRYVVLNRRKNYFGV 672
QY 628 LEDGRIDIVGEAVRGDCWCELAKEVQTVVEI--VLKISEVNKAVEYRKIV---KEL 680
DB 673 TKSQKVDVVKLTGKSKHTPPFTIKELFYSLDLILSAVQTEDEFESAKLIKSKAIAAGKRL 732
QY 681 BEGKVPTEKLVWKTLKRLBEYTTAEHPVVAARMLSAGYRVSPGDKIGYIV--KGG 738
DB 733 EERGVPLADLAFNWMISKAPSEYKTVPOHIRARLLENAREYKGDIIISYKVMNKTGV 792
QY 739 RISQRAWPFYMWDPQSIDVTYYVD 763
DB 793 KPVEMA-----QAGEVDTSKYLE 810

Arch completed: April 23, 2003, 07:46:23
Arch time : 255 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:42:07 ; Search time 24 Seconds
(without alignments)
3216.498 Million cell updates/sec

Title: US-10-034-849-2

Sequence: 1 MTEVVFVTVDSYEVVQKEP.....LKASATGQKTLFPDLAKSR 803

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : PIR 73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	3.5	803	2 B56277	DNA-directed DNA p
2	14	1.7	781	2 UC7382	DNA-directed DNA p
3	14	1.7	784	2 E72515	Probable DNA-direc
4	12	1.5	875	2 JCS186	DNA-directed DNA p
5	12	1.5	1012	2 T41940	DNA polymerase - h
6	11	1.4	1008	2 S55603	DNA polymerase rep
7	11	1.4	1009	1 DJBEK2	DNA-directed DNA p
8	11	1.4	1009	2 T42925	DNA-directed DNA p
9	10	1.2	781	2 A69312	DNA polymerase Bl
10	10	1.2	1015	1 DJBE2L	DNA-directed DNA p
11	10	1.2	1056	1 WMA012	DNA-directed DNA p
12	10	1.2	1056	1 WMA012	DNA-directed DNA p
13	10	1.2	1061	1 DJAD12	DNA-directed DNA p
14	10	1.2	1855	2 E87278	DNA-directed DNA p
15	9	1.1	534	2 E87278	DNA polymerase - m
16	9	1.1	764	2 S75407	sensor histidine k
17	9	1.1	964	2 S25855	Probable DNA-direc
18	9	1.1	1038	2 JCS757	DNA-directed DNA p
19	9	1.1	1038	2 T18222	DNA polymerase del
20	9	1.1	1080	2 FC4394	DNA-directed DNA p
21	9	1.1	1081	2 T20698	hypothetical prote
22	9	1.1	1194	1 DJBE28	DNA-directed DNA p
23	9	1.1	1504	2 A33602	DNA-directed DNA p
24	8	1.0	148	2 S22088	DNA-directed DNA p
25	8	1.0	229	2 T20848	hypothetical prote
26	8	1.0	324	2 A99284	hypothetical prote
27	8	1.0	365	2 AB1828	glycolate oxidase
28	8	1.0	414	2 T33878	hypothetical prote
29	8	1.0	419	2 G86976	Probable flavoprot

30	8	1.0	443	2 D72383	NADH oxidase - The
31	8	1.0	459	2 A82675	heat shock protein
32	8	1.0	514	2 E83827	hypothetical prote
33	8	1.0	520	2 C81273	hypothetical prote
34	8	1.0	521	2 S49439	glucose-1-phosphat
35	8	1.0	563	2 E87086	probable regulator
36	8	1.0	729	2 C89876	hypothetical prote
37	8	1.0	783	1 JDE822	DNA-directed DNA p
38	8	1.0	783	2 H90636	DNA polymerase II
39	8	1.0	783	2 H85487	DNA polymerase II
40	8	1.0	783	2 AE0514	DNA polymerase II
41	8	1.0	787	2 G83410	DNA polymerase II
42	8	1.0	787	2 E82227	DNA polymerase II
43	8	1.0	789	2 AE0064	DNA-directed DNA p
44	8	1.0	844	2 T31321	DNA-directed DNA p
45	8	1.0	879	2 A56277	DNA-directed DNA p
46	8	1.0	882	2 S23019	DNA-directed DNA p
47	8	1.0	882	2 F90201	DNA-directed DNA p
48	8	1.0	889	2 T11742	DNA polymerase I (
49	8	1.0	913	1 A42543	egg sperm receptor
50	8	1.0	913	1 B42543	DNA-directed DNA p
51	8	1.0	913	2 T11765	DNA-directed DNA p
52	8	1.0	959	2 F72763	Probable DNA-direc
53	8	1.0	1012	1 DJBE65	DNA-directed DNA p
54	8	1.0	1012	2 T43998	DNA polymerase (im
55	8	1.0	1012	2 T44185	probable DNA-direc
56	8	1.0	1013	2 TQ1920	DNA-directed DNA p
57	8	1.0	1014	2 T30431	DNA-directed DNA p
58	8	1.0	1026	2 T03108	DNA-directed DNA p
59	8	1.0	1084	1 S19661	DNA-directed DNA p
60	8	1.0	1086	2 T43266	DNA-directed DNA p
61	8	1.0	1086	2 T40242	DNA polymerase del
62	8	1.0	1088	2 T05731	DNA-directed DNA p
63	8	1.0	1094	2 S22573	DNA-directed DNA p
64	8	1.0	1105	1 S40243	DNA-directed DNA p
65	8	1.0	1106	1 A39299	DNA-directed DNA p
66	8	1.0	1107	1 A41618	DNA-directed DNA p
67	8	1.0	1220	1 DJBEC3	DNA-directed DNA p
68	8	1.0	1220	2 T42573	DNA-directed DNA p
69	8	1.0	1242	1 DJBEC1	DNA-directed DNA p
70	8	1.0	1243	2 S07278	tail fiber protein
71	8	1.0	1339	1 S20052	DNA-directed DNA p
72	8	1.0	1339	2 JCS508	DNA-directed DNA p
73	8	1.0	1341	2 S09579	DNA-directed DNA p
74	8	1.0	1405	1 DJZPA	tail fiber protein
75	8	1.0	1462	1 DJHUAC	DNA-directed DNA p
76	8	1.0	1465	2 S45628	DNA-directed DNA p
77	8	1.0	2199	2 T40008	DNA-directed DNA p
78	8	1.0	2272	2 T18572	Cdc20p - fission y
79	7	0.9	42	2 T01991	gag, pol and env p
80	7	0.9	74	2 T47376	hypothetical prote
81	7	0.9	76	2 A72318	hypothetical prote
82	7	0.9	85	2 E83111	hypothetical prote
83	7	0.9	94	2 B64550	hypothetical prote
84	7	0.9	94	2 F71959	hypothetical prote
85	7	0.9	107	2 A29927	ribosomal protein
86	7	0.9	107	2 E70569	hypothetical prote
87	7	0.9	108	2 T50597	hypothetical prote
88	7	0.9	114	2 D98006	hypothetical prote
89	7	0.9	124	2 D90330	hypothetical prote
90	7	0.9	126	2 S52255	hypothetical prote
91	7	0.9	128	2 PN0647	copper resistance
92	7	0.9	130	2 G70799	epsN protein - Vib
93	7	0.9	136	2 F30593	hypothetical prote
94	7	0.9	144	2 G81910	outer membrane pro
95	7	0.9	154	2 A69267	transposase homolo
96	7	0.9	157	1 CYP277	heat shock protein
97	7	0.9	157	2 T40553	hypothetical prote
98	7	0.9	159	2 T32043	hypothetical prote
99	7	0.9	164	2 AD1229	B. subtilis YabB p
100	7	0.9	169	2 D81797	conserved hypothet

ALIGNMENTS

RESULT 1

B56277
DNA-directed DNA polymerase (EC 2.7.7.7) II - Pyrodicticum occultum
C:Species: Pyrodicticum occultum
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000
C:Accession: B56277
R:Uemori, T.; Ishino, Y.; Doi, H.; Kato, I.
J:Biochem Biophys Res Commun 177, 2164-2177, 1995
A:Title: The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-like DNA polymerases
A:Reference number: A56277; MUID:95238290; PMID:7721707
A:Accession: B56277
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-803 <DEM>
A:Cross-references: GB:D36574; NID:9807829; PIDN:BA07580.1; PID:9807830
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: nucleotidyltransferase

Query Match 3.5%; Score 28; DB 2; Length 803;
Best Local Similarity 100.0%; Pred. No. 2,6e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 GLEEDGRIDVGFENVRGDMCLAKEVQ 653
DB 628 GLEEDGRIDVGFENVRGDMCLAKEVQ 655

RESULT 2

JC7382
DNA-directed DNA polymerase (EC 2.7.7.7) B3 - Sulfurisphaera ohwakuensis
A:Alternate names: DNA polymerase alpha, beta, gamma; DNA polymerase I, II, III
C:Species: Sulfurisphaera ohwakuensis
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
C:Accession: JC7382
R:Imai, T.; Kurosawa, N.; Itch, Y.H.; Kimura, N.; Horinuchi, T.
DNA Res. 7, 243-251, 2000
A:Title: Sequence analysis of three family B DNA polymerases from the thermocacidophilic
A:Reference number: JC7380
A:Accession: JC7382
A:Molecule type: DNA
A:Residues: 1-781 <IMA>
A:Cross-references: DDBJ:AB032376
C:Comment: This enzyme has both 3'-5' exonuclease and polymerase activities, and plays a
C:Genetics:
A:Gene: B3
C:Keywords: DNA replication; exonuclease; metal binding; nucleotidyltransferase

Query Match 1.7%; Score 14; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 3,5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 633 IDIVGFENVRGDMC 646
DB 617 IDIVGFENVRGDMC 630

RESULT 3

E72515
Probable DNA-directed DNA polymerase APE2098 - Aeropyrum pernix (strain KI)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: E72515
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: E72515
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-784 <KAW>
A:Cross-references: DDBJ:AF000063; NID:95105654; PIDN:BA081109.1; PID:95105797
A:Experimental source: strain KI
C:Genetics:
A:Gene: APE2098
C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 1.7%; Score 14; DB 2; Length 784;
Best Local Similarity 100.0%; Pred. No. 3,5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 449 KYNVGPDLVVRPGE 462
DB 433 KYNVGPDLVVRPGE 446

RESULT 4

JC5186
DNA-directed DNA polymerase (EC 2.7.7.7) - Sulfolobus acidocaldarius
C:Species: Sulfolobus acidocaldarius
C:Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 18-Jun-1999
C:Accession: JC5186; PC4255
R:Drumstovill, N.; Pokholok, D.; Lottepeich, F.; Prangishvili, D.; Rechin, V.
Gene 177, 271-273, 1996
A:Title: The DNA polymerase-encoding gene from a thermocacidophilic archaeon Sulfolobus
A:Reference number: JC5186; MUID:97080536; PMID:8921861
A:Accession: JC5186
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-875 <DAT1>
A:Cross-references: GB:U3846; NID:91685081; PIDN:AAC44598.1; PID:91685082
A:Accession: PC4255
A:Status: preliminary
A:Molecule type: protein
A:Residues: 713-726; 777-790 <DAT2>
C:Comment: This enzyme belongs to family B of DNA polymerase, and has proof reading acti
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: nucleotidyltransferase
F:223-235/Region: exonuclease signature
F:286-300/Region: exonuclease signature
F:402-414/Region: exonuclease signature

Query Match 1.5%; Score 12; DB 2; Length 875;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 572 LKVIYGDPSLF 583
DB 646 LKVIYGDPSLF 657

RESULT 5

T41940
DNA polymerase - human herpesvirus 7 (strain J1)
C:Species: human herpesvirus 7
A:Variety: strain J1
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T41940
R:Nicholas, J.
Submitted to the EMBL Data Library, December 1995
A:Description: Determination and analysis of the complete nucleotide sequence of human
A:Reference number: 222022
A:Accession: T41940
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1012 <NIC>
A:Cross-references: EMBL:U03400; PIDN:AAC54700.1
A:Experimental source: strain J1
C:Genetics:
A:Note: U38
C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 1.5%; Score 12; DB 2; Length 1012;

Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDITSLFV 585
DB 735 VIYGDITSLFV 746

RESULT 6

S55603
DNA polymerase replicative chain - equine herpesvirus 2

C/Species: equine herpesvirus 2
C/Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 18-Jun-1999

C/Accession: S55603
R/RefSeq: E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.

J. Mol. Biol. 249, 520-528, 1995

A/Title: The DNA sequence of equine herpesvirus 2.

A/Reference number: S55594; MUID:95302501; PMID:7783207

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Accession: S55603
A/Molecule type: DNA

A/Residues: 1-1008 <TEL>

A/Cross-references: GB:U20824; NID:G695172; PIDN:AA013796.1; PID:G695181

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 1.4%; Score 11; DB 2; Length 1008;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 573 KVIYGDITSLF 583
DB 744 KVIYGDITSLF 754

RESULT 7

DJBEM2
DNA-directed DNA polymerase (EC 2.7.7.7) - saimiriine herpesvirus 1 (strain 11)

C/Species: saimiriine herpesvirus 1
A/Note: host Saimiri sciureus (common squirrel monkey)

C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 11-Jun-1999

C/Accession: I36806
R/RefSeq: J.

A/Title: Submitted to the EMBL Data Library, January 1992

A/Description: Primary structure of the herpesvirus saimiri genome.

A/Reference number: A36806

A/Accession: I36806
A/Molecule type: DNA

A/Residues: 1-1009 <ALB>

A/Cross-references: GB:X64346; NID:G60320; PIDN:CAA5632.1; PID:G60330

R/RefSeq: J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; W.

J. Virol. 66, 5047-5058, 1992

A/Title: Primary structure of the herpesvirus saimiri genome.

A/Reference number: A37309; MUID:9233688; PMID:131287

A/Contents: annotation; possible protein-coding frames

A/Note: neither amino acid nor nucleotide sequence is given

C/Genetics:

A/Gene: 9

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

QY 574 VIYGDITSLFV 584
DB 737 VIYGDITSLFV 747

RESULT 8

T42925
DNA-directed DNA polymerase (EC 2.7.7.7) - ateline herpesvirus 3 (strain 73)

C/Species: ateline herpesvirus 3
A/Variety: strain 73
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C/Accession: T42925

R/RefSeq: J.C.; Fleckenstein, B.

submitted to the EMBL Data Library, August 1998

A/Description: Primary structure of the herpesvirus ateles genome.

A/Reference number: Z22274

A/Accession: T42925

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1009 <ALB>

A/Cross-references: EMBL:AF083424; PIDN:AA095533.1

A/Experimental source: strain 73

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: nucleotidyltransferase

Query Match 1.4%; Score 11; DB 2; Length 1009;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDITSLFV 584
DB 737 VIYGDITSLFV 747

RESULT 9

A69312
DNA polymerase B1 (polB) homolog - Archaeoglobus fulgidus

C/Species: Archaeoglobus fulgidus

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C/Accession: A69312

R/RefSeq: H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A/Authors: Ufferkamp, T.; Cotton, M.D.; Spriggs, T.; Arriach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A/Reference number: A69250; MUID:98049343; PMID:9389475

A/Accession: A69312

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-781 <ALB>

A/Cross-references: GB:AE001070; GB:AE000782; NID:G268993; PIDN:AA090741.1; PID:G265013

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: herpesvirus DNA-directed DNA polymerase

QY 239 YDPDIIVGVN 248
DB 218 YDPDIIVGVN 227

RESULT 10

DJBEM2
DNA-directed DNA polymerase (EC 2.7.7.7) - human herpesvirus 4 (strain B95-8)

C/Species: human herpesvirus 4, Epstein-Barr virus

C/Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 11-Jun-1999

C/Accession: A00713; S33054

R/RefSeq: A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.

Mol. Biol. Med. 1, 21-45, 1983

A/Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus

A/Reference number: A93065; MUID:8503573; PMID:6092825

A/Accession: A00713

A/Molecule type: DNA

A/Residues: 1-1015 <BAN>

A/Cross-references: EMBL:V01555; NID:G59074; PIDN:CAA24805.1; PID:G1334913

R/RefSeq: A.T.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H

Nature 310, 207-211, 1984

A/Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.

A:Reference number: A03794; MUID:84270667; PMID:6087149
A:Content: annotation; protein coding region
C:Species: herpesvirus DNA-directed DNA polymerase
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.2%; Score 10; DB 1; Length 1015;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 YGDTDSLFT 583
DB 751 YGDTDSLFT 760

RESULT 11
WMA012
DNA-directed DNA polymerase (EC 2.7.7.7) - human adenovirus 2
C:Species: Mastadenovirus h2 (human adenovirus 2)
A:Note: host Homo sapiens (man)
C>Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 16-Feb-1997
C:Accession: A92351; A92352; A00711
C:Reference number: A92351; A92352; A00711
A:Title: Nucleotide sequences from the adenovirus-2 genome.
A:Reference number: A92351; MUID:83056843; PMID:7142161
A:Accession: A92351
A:Molecule type: DNA
A:Residues: 1-1056 <GIN>
R:Abstract: P.; Akusjarvi, G.; Pettersson, M.; Pettersson, U.
U. Biol. Chem. 257, 13492-13498, 1982
A:Title: DNA sequence analysis of the region encoding the terminal protein and the hypoc
A:Reference number: A92352; MUID:83056844; PMID:7142162
A:Accession: A92352
A:Molecule type: DNA
A:Residues: 1-1056 <ALE>
C:Superfamily: adenovirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.2%; Score 10; DB 1; Length 1056;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGDTDSLFT 585
DB 868 YGDTDSLFT 877

RESULT 12
BADA01
DNA-directed DNA polymerase (EC 2.7.7.7) - human adenovirus 5
C:Species: Mastadenovirus h5 (human adenovirus 5)
A:Note: host Homo sapiens (man)
C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 11-Jun-1999
C:Accession: A00712
C:Reference number: A00712
R:Decker, B.M.M.; van Ormondt, H.
Gene 27, 115-120, 1984
A:Title: The nucleotide sequence of fragment HindIII-C of human adenovirus type 5 DNA (a
A:Reference number: A91508; MUID:84183604; PMID:6325298
A:Accession: A00712
A:Molecule type: DNA
A:Residues: 1-1056 <DEK>
A:Cross-references: EMBL:X02996; NID:958484; PIDN:CAA26749.1; PID:958495
C:Superfamily: adenovirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.2%; Score 10; DB 1; Length 1056;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGDTDSLFT 585
DB 868 YGDTDSLFT 877

Query Match 1.2%; Score 10; DB 1; Length 1056;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGDTDSLFT 585
DB 868 YGDTDSLFT 877

RESULT 13
DADA12
DNA-directed DNA polymerase (EC 2.7.7.7) - human adenovirus 12
C:Species: Mastadenovirus h12 (human adenovirus 12)
C>Date: 30-Jun-1989 #sequence_revision 17-May-1996 #text_change 11-Jun-1999
C:Accession: S33933; A25770
R:Spiegel, J.
submitted to the EMBL Data Library, June 1993
A:Reference number: S33928
A:Accession: S33933
A:Molecule type: DNA
A:Residues: 1-1061 <SPR>
A:Cross-references: EMBL:X73487; NID:931361; PIDN:CAA51882.1; PID:931367
R:Shu, L.; Hong, J.S.; Wei, Y.F.; Engler, J.A.
Gene 46, 187-195, 1986
A:Title: Nucleotide sequence of the genes encoded in early region 2b of human adenovirus
A:Reference number: A91557; MUID:87106854; PMID:3803925
A:Accession: A25770
A:Molecule type: DNA
A:Residues: 931, 'S', 33-163, 'T', 163-180, 'YN', 183-460, 'T', 462-574, 'F', 576-891, 'T', 893-102
C:Comment: This gene is located on the E2b region of the genome.
C:Superfamily: adenovirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; early protein; nucleotidyltransferase

Query Match 1.2%; Score 10; DB 1; Length 1061;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGDTDSLFT 585
DB 874 YGDTDSLFT 883

RESULT 14
S41649
DNA polymerase - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 20-May-1994 #sequence_revision 26-May-1995 #text_change 09-Jun-2000
C:Accession: S41649
R:Whittle, J.H.; Kilbey, B.J.; de Vries, E.; Goman, M.; Alano, P.; Cheesman, S.; McAlleese,
Nucleic Acids Res 21, 3643-3646, 1993
A:Title: The gene encoding DNA polymerase alpha from Plasmodium falciparum.
A:Reference number: S41649; MUID:93576482; PMID:8367280
A:Accession: S41649
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1855 <WHI>
A:Cross-references: EMBL:L18785
C:Superfamily: Plasmodium falciparum DNA polymerase

Query Match 1.2%; Score 10; DB 2; Length 1855;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 572 LKVIYGDTS 581
DB 1237 LKVIYGDTS 1246

RESULT 15
E87278
sensor histidine kinase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: E87278
R:Nietman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, U.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gilm, M.L.; Hafe, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venturi, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87278

A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-534 <SNO>
A:Cross-references: GB:AE005673; NID:g13421367; PIDN:AAK22225.1; GSPDB:GN00148
A:Gene: CC0238

Query Match
Best Local Similarity 1.1%; Score 9; DB 2; Length 534;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 IAENGERV 35
Db 181 IAENGERV 189

RESULT 16

75407
A:Status: DNA-directed DNA polymerase (EC 2.7.7.7) - Sulfolobus solfataricus
A:Alternate names: protein c04041
C:Species: Sulfolobus solfataricus
C>Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C:Accession: S75407
R:Seisen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A>Title: Organizational characteristics and information content of an archaeal genome: 1
A:Reference number: S73076; MUID:97055432; PMID:8899719
A:Accession: S75407
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-764 <SEN>
A:Cross-references: EMBL:Y08257; NID:g1707772; PID:g1707813
A:Experimental source: strain P2
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996.
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: nucleotidyltransferase

Query Match
Best Local Similarity 1.1%; Score 9; DB 2; Length 764;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 728 KIGYIVKG 736
Db 691 KIGYIVKG 699

RESULT 17

5855
A:Status: DNA-directed DNA polymerase (EC 2.7.7.7) - Choriostoneura biennis poxvirus
C:Species: Choriostoneura biennis poxvirus
C>Date: 08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999
C:Accession: S25855
R:Mustafa, A.; Yuen, L.
DNA Seq. 2, 39-45, 1991
A>Title: Identification and sequencing of the Choriostoneura biennis entomopoxvirus DNA
A:Reference number: S25855; MUID:92199242; PMID:1840516
A:Accession: S25855
A:Molecule type: DNA
A:Residues: 1-964 <MUS>
A:Cross-references: EMBL:X57314; NID:958845; PIDN:CAA40566.1; PID:958846
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: DNA binding; nucleotidyltransferase

Query Match
Best Local Similarity 1.1%; Score 9; DB 2; Length 964;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 575 IYGDTSLF 583
Db 750 IYGDTSLF 758

RESULT 18

JC5757

DNA-directed DNA polymerase (EC 2.7.7.7) III - yeast (Candida albicans)
A:Alternate names: DNA polymerase III
C:Species: Candida albicans
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Dec-2000
C:Accession: JC5757; S60677
R:Nolan, T.; Rosemond, J.
Gene 183, 159-165, 1996
A>Title: Isolation and molecular characterization of the POL3 gene from Candida albicans
A:Reference number: JC5757; MUID:97149294; PMID:8996102
A:Accession: JC5757
A:Molecule type: DNA
A:Residues: 1-1038 <NOL>
A:Cross-references: EMBL:X88804
A>Note: Submitted to the EMBL Data Library, June 1995
C:Comment: This enzyme plays a role in DNA replication.
C:Gene: POL3
A:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: nucleotidyltransferase; zinc finger
F:942-961/Region: zinc finger CCCC motif
F:992-1010/Region: zinc finger CCCC motif

Query Match
Best Local Similarity 1.1%; Score 9; DB 1; Length 1038;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 573 KVIYGDTS 581
Db 685 KVIYGDTS 693

RESULT 19

T18222
A:Status: DNA polymerase delta chain - yeast (Candida albicans)
C:Species: Candida albicans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T18222
R:Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, November 1998
A:Reference number: Z18831
A:Accession: T18222
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1038 <BAR>
A:Cross-references: EMBL:AL033396; PIDN:CAA21949.1
A:Gene: C
A:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match
Best Local Similarity 1.1%; Score 9; DB 2; Length 1038;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 573 KVIYGDTS 581
Db 685 KVIYGDTS 693

RESULT 20

PC4394
A:Status: DNA-directed DNA polymerase (EC 2.7.7.7) - Ovine adenovirus OAV287 (fragment)
C:Species: Ovine adenovirus OAV287
C>Date: 10-Nov-1997 #sequence_revision 10-Nov-1997 #text_change 03-Nov-2000
C:Accession: PC4394
R:Watt, S.; Brooks, D.E.; Boyle, D.B.; Both, G.W.
Gene 177, 35-41, 1996

A>Title: Nucleotide sequence of ovine adenovirus tripartite leader sequence and homolog
A:Reference number: JC5648; MUID:97080497; PMID:8921842
A:Accession: PC4394
A:Molecule type: DNA
A:Residues: 1-1080 <VRA>
A:Cross-references: GB:UJ1557; NID:g1117828; PIDN:AAK55557.1; PID:g1117820

C/Comment: This enzyme is targeted to the nucleus by interaction with the terminal prote

C/Superfamily: adenovirus DNA-directed DNA polymerase

C/Keywords: nucleotidyltransferase

Query Match 1.1%; Score 9; DB 2; Length 1080;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 575 IYGDTSLPF 583
DB 882 IYGDTSLPF 890

RESULT 21
T20698
hypothetical protein F10C2.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T20698
R/White, S.
Submitted to the EMBL Data Library, November 1996
Reference number: Z19311
Accession: T20698
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1081 <N1>
A/Cross-references: EMBL:Z81497; PIDN:CA804077.1; GSPDB:GN00023; CESP:F10C2.4
A/Experimental source: clone F10C2
C/Genetics:
A/Gene: CESP:F10C2.4
A/Map position: 5
A/Intons: 29/3; 112/3; 216/3; 281/2; 524/3; 862/3; 1012/2
C/Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 1.1%; Score 9; DB 2; Length 1081;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 573 KYIYGTDS 581
DB 721 KYIYGTDS 729

RESULT 22
DJB228
DNA-directed DNA polymerase (EC 2.7.7.7) - human herpesvirus 3
C/Species: human herpesvirus 3, varicella-zoster virus
C/Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 11-Jun-1999
C/Accession: B27214
A/Title: The complete DNA sequence of varicella-zoster virus.
A/Reference number: A27345; M01D:86306657; PMID:3018124
A/Accession: B27214
A/Molecule type: DNA
A/Residues: 1-1194 <DAN>
A/Cross-references: EMBL:X04370; NID:G59989; PIDN:CAA27911.1; PID:G60017
C/Genetics:
A/Gene: 28
C/Superfamily: herpesvirus DNA-directed DNA polymerase
C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.1%; Score 9; DB 1; Length 1194;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 573 KYIYGTDS 581
DB 846 KYIYGTDS 854

RESULT 23
A33602

DNA-directed DNA polymerase (EC 2.7.7.7) REV3 - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein p255; protein YPL167C
C/Species: Saccharomyces cerevisiae
C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 20-Jun-2000
C/Accession: A33602; S65178; S69432
R/Morrison, A.; Christensen, R.B.; Alley, J.; Beck, A.K.; Bernatine, E.G.; Lemont, J.F.
J. Bacteriol. 171, 5659-5667, 1989
A/Title: REV3, a Saccharomyces cerevisiae gene whose function is required for induced m
A/Reference number: A33602; M01D:90008808; PMID:2676986
A/Accession: A33602
A/Molecule type: DNA
A/Residues: 1-1504 <NOR>
A/Cross-references: GB:M29683; NID:G172386; PIDN:AAA34968.1; PID:G172387
A/Experimental source: strain S288C
R/Purnelle, B.; Coster, F.; Goffeau, A.
Submitted to the Protein Sequence Database, May 1996
A/Reference number: S65154
A/Accession: S65178
A/Molecule type: DNA
A/Residues: 1-1504 <PUR>
A/Cross-references: EMBL:Z73523; NID:G1370352; PIDN:CAA97873.1; PID:G1370353; MIPS:YPL16
A/Experimental source: strain S288C (A5972)
R/Purnelle, B.; Comblez, S.; Coster, F.; Naveau, F.; Goffeau, A.
Submitted to the EMBL Data Library, March 1996
A/Description: The sequence of 55 kb on the left arm of yeast chromosome XVI identifies
A/Reference number: S69428
A/Accession: S69432
A/Molecule type: DNA
A/Residues: 1-1504 <PRM>
A/Cross-references: EMBL:X96770; NID:G1403537; PIDN:CAA65554.1; PID:G1403542
C/Genetics:
A/Gene: SGD:REV3; PS01
A/Cross-references: SGD:S0006088; MIPS:YPL167C
A/Map position: 16L
C/Superfamily: yeast DNA-directed DNA polymerase REV3
C/Keywords: DNA binding; nucleotidyltransferase

Query Match 1.1%; Score 9; DB 2; Length 1504;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 576 YGPTDSLFPV 584
DB 1140 YGPTDSLFPV 1148

RESULT 24
S22088
DNA-directed DNA polymerase (EC 2.7.7.7) delta catalytic chain - slime mold (Physarum pc
C/Species: Physarum polycephalum
C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 18-Jun-1999
C/Accession: S22088; S65029
R/Holler, E.
Submitted to the EMBL Data Library, May 1992
A/Reference number: S22088
A/Accession: S22088
A/Molecule type: mRNA
A/Residues: 1-148 <HOL>
A/Cross-references: EMBL:X66049; NID:G3192; PIDN:CAA46847.1; PID:G3193
A/Experimental source: strain M3C VIII
R/Achhammer, G.; Winkler, A.; Angerer, B.; Holler, E.
Curr. Genet. 28, 534-545, 1995
A/Title: DNA polymerase delta of Physarum polycephalum.
A/Reference number: S65029; M01D:96132109; PMID:8593684
A/Accession: S65029
A/Molecule type: mRNA
A/Residues: 1-142 <ACH>
A/Cross-references: EMBL:X66049
C/Genetics:
A/Gene: POL3
C/Superfamily: herpesvirus DNA-directed DNA polymerase
C/Keywords: DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDIDS 581
Db 141 VIYGDIDS 148

RESULT 25

T20848
hypothetical protein F13E9.10 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C/Accession: T20848
R/McMurray, A.
Submitted to the EMBL Data Library, February 1996

A/Accession: T20848
Reference number: 219332

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-229 <WIL>

A/Cross-references: EMBL:Z69383; PIDN:CAA93413.1; GSPDB:GN00022; CESP:F13E9.10

A/Experimental source: clone F13E9

C/Genetics:
A/Gene: CESP:F13E9.10
A/Map position: 4
A/Intons: 65/3; 127/3

C/Superfamily: *Caenorhabditis elegans* hypothetical protein F13E9.10

Query Match 1.0%; Score 8; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 700 LEEYTTA 707
Db 165 LEEYTTA 172

RESULT 26

A99284
hypothetical protein SSO1290 [imported] - *Sulfolobus solfataricus*

C/Species: *Sulfolobus solfataricus*
C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C/Accession: A99284
She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-

g, R.A.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.

Submitted to Genbank, April 2001
A/Description: *Sulfolobus solfataricus* complete genome.

A/Reference number: A99139

A/Accession: A99284
A/Status: preliminary

A/Molecule type: DNA
A/Residues: 1-324 <KUR>

A/Cross-references: GB:AE006641; NID:913814490; PIDN:AAK41528.1; GSPDB:GN00155

C/Genetics:
A/Gene: SSO1290

Query Match 1.0%; Score 8; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 ERLLELRK 499
Db 43 ERLLELRK 50

RESULT 27

ABI828
glycolate oxidase [imported] - *Nostoc* sp. (strain PCC 7120)

C/Species: *Nostoc* sp.

A/Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C/Accession: ABI828
R/Xaneco, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*

A/Reference number: ABI807; MUID:21595285; PMID:1175840

A/Accession: ABI828
A/Status: preliminary

A/Molecule type: DNA
A/Residues: 1-365 <KUR>

A/Cross-references: GB:BA000019; PIDN:BA877694.1; PID:917135148; GSPDB:GN00179

A/Experimental source: strain PCC 7120

C/Genetics:
A/Gene: all0170
C/Superfamily: (S)-2-hydroxy-acid oxidase; (S)-2-hydroxy-acid oxidase homology

Query Match 1.0%; Score 8; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 309 KSLEEVAE 316
Db 115 KSLEEVAE 122

RESULT 28

T33878
hypothetical protein Y14H12B.2 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C/Accession: T33878
R/Rohlfing, T.; Wohldman, P.; Courtney, L.; Maupin, R.

Submitted to the EMBL Data Library, February 1999

A/Description: The sequence of C. elegans cosmid Y14H12B.

A/Reference number: Z21430

A/Accession: T33878
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
A/Residues: 1-414 <ROH>

A/Cross-references: EMBL:AF125447; PIDN:AD12807.1; GSPDB:GN00020; CESP:Y14H12B.2

A/Experimental source: strain Bristol NZ; clone Y14H12B

C/Genetics:
A/Gene: CESP:Y14H12B.2
A/Map position: 2
A/Intons: 105/2; 272/2; 394/3

Query Match 1.0%; Score 8; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 796 DFLAKSK 803
Db 29 DFLAKSK 36

RESULT 29

G86976
probable flavoprotein [imported] - *Mycobacterium leprae*

C/Species: *Mycobacterium leprae*
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C/Accession: G86976
R/Cole, S.T.; Eiglmeyer, K.; Parhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hc

R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holroyd,

eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001

A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq

A/Title: Massive gene decay in the leprosy bacillus.

A/Reference number: AB6909; MUID:21128732; PMID:11234002

A/Accession: G86976
A/Status: preliminary

A/Molecule type: DNA

A:Residues: 1-419 <STO>
A:Cross-references: GB:AL450380; NID:G13092747; PIDN:CACG30051.1; GSPDB:GN00147
C:Genetics:
A:Gene: dfr
C:Superfamily: pantothenate metabolism flavoprotein dfr

Query Match 1.0%; Score 8; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 419 RGATVLP 426
DB 140 RGATVLP 147

RESULT 30

D72383
NADH oxidase - Thermotoga maritima (strain MSB)

C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

A:Accession: D72383
A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <ARN>
A:Cross-references: GB:AE001718; GB:AE000512; NID:G4980881; PIDN:ADJ35465.1; PID:G498088

C:Genetics:
A:Gene: TM0379

C:Superfamily: NADH peroxidase; dihydroliipoamide dehydrogenase homology
F:6-429/Domain: dihydroliipoamide dehydrogenase homology <DID>

Query Match 1.0%; Score 8; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 ARKGLKV 574
DB 245 ARKGLKV 252

RESULT 31
A82675
Heat shock protein Xf1485 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000

A:Accession: A82675
A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-459 <SIM>
A:Cross-references: GB:AE003978; GB:AE003849; NID:G9106510; PIDN:AMP84294.1; GSPDB:GN001

C:Genetics:
A:Gene: dfr

C:Superfamily: pantothenate metabolism flavoprotein dfr

Query Match 1.0%; Score 8; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 419 RGATVLP 426
DB 140 RGATVLP 147

RESULT 32

D72383
NADH oxidase - Thermotoga maritima (strain MSB)

C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

A:Accession: D72383
A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <ARN>
A:Cross-references: GB:AE001718; GB:AE000512; NID:G4980881; PIDN:ADJ35465.1; PID:G498088

C:Genetics:
A:Gene: TM0379

C:Superfamily: NADH peroxidase; dihydroliipoamide dehydrogenase homology
F:6-429/Domain: dihydroliipoamide dehydrogenase homology <DID>

Query Match 1.0%; Score 8; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 ARKGLKV 574
DB 245 ARKGLKV 252

RESULT 33

D72383
NADH oxidase - Thermotoga maritima (strain MSB)

C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

A:Accession: D72383
A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <ARN>
A:Cross-references: GB:AE001718; GB:AE000512; NID:G4980881; PIDN:ADJ35465.1; PID:G498088

C:Genetics:
A:Gene: TM0379

C:Superfamily: NADH peroxidase; dihydroliipoamide dehydrogenase homology
F:6-429/Domain: dihydroliipoamide dehydrogenase homology <DID>

Query Match 1.0%; Score 8; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 ARKGLKV 574
DB 245 ARKGLKV 252

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasat,
M.; Tounheto, M.H.; Valiada, H.; Van Silve, M.A.; Verjovskij, Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Status: annotation
C:Genetics:
A:Gene: Xf1485
C:Superfamily: heat shock protein hsp10; FtsH/Sec18/CDC48-type ATP-binding domain homolo.

Query Match 1.0%; Score 8; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 489 TYLERLE 496
DB 413 TYLERLE 420

RESULT 32

E83827
Hypothetical protein BH1421 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

A:Accession: E83827
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-514 <STO>
A:Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BAE05140.1; GSPDB:GN0C

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-514 <STO>
A:Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BAE05140.1; GSPDB:GN0C

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-514 <STO>
A:Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BAE05140.1; GSPDB:GN0C

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-514 <STO>
A:Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BAE05140.1; GSPDB:GN0C

Query Match 1.0%; Score 8; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 518 LDEROKAL 525
DB 482 LDEROKAL 489

RESULT 33

C81273
Hypothetical protein Cj1302 [imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

A:Accession: C81273
A:Reference number: A81250; MUID:20150912; PMID:110688204

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-520 <PAR>
A:Cross-references: GB:AL139078; GB:AL111168; NID:G6968723; PIDN:CAB73729.1; PID:G696873

C:Genetics:
A:Gene: Cj1302

C:Superfamily: heat shock protein hsp10; FtsH/Sec18/CDC48-type ATP-binding domain homolo.

Query Match 1.0%; Score 8; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 ERLERLK 499
DB 97 ERLERLK 104

RESULT 34

S49439

glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) - maize

C:Species: Zea mays (maize)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999

C:Accession: S49439

R:Smith-White, B.J.; Gilroy, M.J.; Gilmore, V.; Preiss, J.; Hannah, L.C.

A:Description: The EMBL Data Library, October 1994

A:Reference number: S49439

A:Accession: S49439

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1521 <SMT>

A:Cross-references: EMBL:Z28111; NID:G558364; PID:G558365

A:Function:

Description: catalyzes the formation of ADPGlucose and pyrophosphate from alpha-D-glucose-1-phosphate and ATP

C:Superfamily: glucose-1-phosphate adenylyltransferase

C:Keywords: glycogen/starch biosynthesis; nucleotidyltransferase

Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 521;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 453 GPDVLRP 460

DB 66 GPDVLRP 73

RESULT 35

E87086

probable regulatory protein [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: E87086

R:Cole, S.T.; Eiglsmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Heam, M.A.; Rutherford, K.M.

A:Authors: Rutherford, K.M.

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: E87086

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1563 <STO>

A:Cross-references: GB:AL450380; NID:G13093298; PID:GAC30370.1; GSPDB:GN00147

C:Genetics:

A:Gene: MLI419

Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 563;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 768 PAARITG 775

DB 286 PAARITG 293

RESULT 36

C89876

hypothetical protein purL [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: C89876

R:Kuroda, M.; Ohba, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shibata, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

A:Reference number: S12820; MUID:91088346; PMID:124684

A:Status: nucleic acid sequence not shown

A:Reference number: A89758; MUID:21111952; PMID:11418146

A:Accession: C89876

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1729 <KUR>

A:Cross-references: GB:BA000018; PID:G13700870; PID:BA042166.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: purL

C:Superfamily: phosphoribosylformylglycinamide synthase component II

Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 729;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 605 EIKLEKRY 612

DB 11 EIKLEKRY 18

RESULT 37

JDBC22

DNA-directed DNA polymerase (EC 2.7.7.7) II - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 01-Mar-2002

C:Accession: S15943; S19262; S40576; J00780; A38840; S19263; S12820; B36236; D64727; S54

R:Wassak, H.; Ishino, Y.; Toh, H.; Nakata, A.; Shingawa, H.

A:Reference number: S15943; MUID:91238699; PMID:2034216

A:Accession: S15943

A:Molecule type: DNA

A:Residues: 1783 <IWA>

A:Cross-references: EMBL:X54847; NID:G42462; PID:CAA38616.1; PID:G581193

A:Experimental source: strain K-12, substrain W3110

A:Accession: S19262

A:Molecule type: protein

A:Residues: 29 <IWA2>

R:Turn, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobu

submitted to the EMBL Data Library, December 1992

A:Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-2

A:Reference number: S40531

A:Accession: S40576

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 'V', '2-783 <YUR>

A:Cross-references: EMBL:D10483; NID:G216434

A:Experimental source: strain K-12

R:Bonner, C.A.; Hays, S.; McEntee, K.; Goodman, M.F.

P.N.C. Natl. Acad. Sci. U.S.A. 87, 7663-7667, 1990

A:Title: DNA polymerase II is encoded by the DNA damage-inducible dna gene of Escherich

A:Reference number: J00780; MUID:91017565; PMID:2217198

A:Accession: J00780

A:Molecule type: DNA

A:Residues: 1-256, 'DD', '259-271', 'G', '273-458 <BON>

A:Cross-references: GB:M37727; NID:G145744; PID:AAA23684.1; PID:G145746

A:Accession: A38840

A:Molecule type: protein

A:Residues: 2-9, 'XQ', '12-21', 'H', '23', 'X', '25-28 <BO2>

R:Chen, H.

submitted to the EMBL Data Library, June 1990

A:Reference number: S19263

A:Accession: S19263

A:Molecule type: DNA

A:Residues: 1-171, 'A', '173-734', 'T', '736-739', 'PGPFTTGRTSDPPATRRG', '763', 'NT', '766-767', 'Y'

A:Cross-references: EMBL:M35311; NID:G147319; PID:AAA24407.1; PID:G147320

Nucleic Acids Res. 18, 7185-7186, 1990

A:Title: Aphidicolin inhibits DNA polymerase II of Escherichia coli, an alpha-like DNA p

A:Reference number: S12820; MUID:91088346; PMID:124684

A:Accession: S12820

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 90-171, 'A', 173-734, 'T', 736-739, 'PGLPTTGLRLTSPDPRATTRGG', 763, 'NT', 766-767, 'A'
A:Cross-references: EMBL:M35371
R:Chen, H.; Sun, Y.; Stark, T.; Beattie, W.; Moses, R.E.
DNA Cell Biol 9, 631-635, 1990
A:Title: Nucleotide sequence and deletion analysis of the *polB* gene of *Escherichia coli*.
A:Reference number: A36236; MUID:91083833; PMID:2261080
A:Accession: B36236
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <BLAT>
A:Cross-references: GB:AE000116; GB:U00996; NID:G1786240; PIND:AA03171.1; PID:G1786246;
Experimental source: strain K-12, substrain MG1655
A:Accession: GJ
A:Accession: S54807
A:Reference number: S54807
A:Molecule type: mRNA
A:Residues: 428-444, 'LIQSTVPRV', 454, 'SMPSREKNIAV', 467-482, 'RVTV', 487-510, 'L', 512-520, 'T'
A:Cross-references: EMBL:X87158; NID:9809066; PID:9809067
A:Note: The source was identified as *Physarum polycephalum*, this is probably a misassignment
A:Note: compared to *Escherichia coli* there is 96% sequence identity at the DNA level and
R:Ackhammer, G.; Winkler, A.; Angerer, B.; Holler, E.
Curr. Genet. 28, 534-545, 1995
A:Title: DNA polymerase delta of *Physarum polycephalum*.
A:Reference number: S65029; MUID:96132109; PMID:8593684
A:Accession: S65030
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 428-444, 'LIQSTVPRV', 454, 'SMPSREKNIAV', 467-482, 'RVTV', 487-510, 'L', 512-520, 'T'
A:Cross-references: EMBL:X87158
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
A:Note: the source was identified as *Physarum polycephalum*; this is probably a misassignment
A:Note: compared to *Escherichia coli* there is 96% sequence identity at the DNA level and
C:Genetics:
A:Gene: *polB*
A:Map position: 2 min
A:Start codon: GTG
A:Function:
A:Pathway: purine metabolism; pyrimidine metabolism
A:Superfamily: *Escherichia coli* DNA-directed DNA polymerase II
A:Keyword: DNA binding; DNA biosynthesis; nucleic acid metabolism
A:Product: DNA-directed DNA polymerase II [status experimental] <MAT>
Query Match 1.0%; Score 8; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 574 VIYGDTS 581
Db 541 VIYGDTS 548
RESULT 39
H90636
DNA polymerase II [imported] - *Escherichia coli* (strain O157:H7, substrain RMD 0509952)
C:Species: *Escherichia coli*
C>Date: 16-Jul-2001 #sequence_revision 16-Jul-2001 #ext_change 03-Aug-2001
C:Accession: H90636
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
Sakawara, N.; Yasunaga, T.; Kuwara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: H90636
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <HAV>
A:Cross-references: GB:BA000007; PIND:BA033467.1; PID:G13359520; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC9064
C:Superfamily: *Escherichia coli* DNA-directed DNA polymerase II
Query Match 1.0%; Score 8; DB 2; Length 783;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 574 VIYGDTS 581
Db 541 VIYGDTS 548
RESULT 39
H85487
DNA polymerase II [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)
C:Species: *Escherichia coli*
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #ext_change 14-Sep-2001
C:Accession: H85487
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85487
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <STO>
A:Cross-references: GB:AE005174; NID:G12512757; PIND:AA054364.1; GSPDB:GN00145; UWGP:200
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: *polB*
A:Superfamily: *Escherichia coli* DNA-directed DNA polymerase II
Query Match 1.0%; Score 8; DB 2; Length 783;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 574 VIYGDTS 581
Db 541 VIYGDTS 548
RESULT 40
AE0514
DNA polymerase II [imported] - *Salmonella enterica* subsp. *enterica* serovar Typh1 (strain
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh1
A:Note: this species has also been called *Salmonella typh1*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #ext_change 27-Nov-2001
C:Accession: AE0514
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
Th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dord, L.; White, N.; Farrar,
S.; Moulé, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov.
A:Reference number: AB0502; PMID:11677608
A:Accession: AE0514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <PAR>
A:Cross-references: GB:AL513382; PIND:CA001253.1; PID:G16501382; GSPDB:GN00176
C:Genetics:
A:Gene: STY0112
C:Superfamily: *Escherichia coli* DNA-directed DNA polymerase II
Query Match 1.0%; Score 8; DB 2; Length 783;

Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
|||||
DB 541 VIYGDTS 548

RESULT 41

DNA polymerase II PA186 [imported] - Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C/Accession: G83410

R/Stover, C.K.; Pham, X.O.; Errin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

Reference number: A82950; MUID:20437337; PMID:10984043

Accession: G83410

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-787 <STO>

A/Cross-references: GB:AE004614; GB:AE004091; NID:G9947868; PIDN:AG05275.1; GSPDB:GN001

A/Experimental source: strain PA01

C/Genetics:

A/Gene: polB; PA186

C/Superfamily: Escherichia coli DNA-directed DNA polymerase II

Query Match 1.0%; Score 8; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
|||||
DB 543 VIYGDTS 550

RESULT 42

DNA polymerase II VC1212 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C/Species: Vibrio cholerae

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C/Accession: E82227

R/Haidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

ardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.

R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Accession: E82227

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-787 <HEI>

A/Cross-references: GB:AE004201; GB:AE003852; NID:G9555689; PIDN:AF94371.1; GSPDB:GN001

C/Experimental source: serogroup O1; strain N16961; biotype El Tor

C/Genetics:

A/Gene: VC1212

A/Map position: 1

C/Superfamily: Escherichia coli DNA-directed DNA polymerase II

Query Match 1.0%; Score 8; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
|||||
DB 544 VIYGDTS 551

RESULT 43

AE0064

DNA-directed DNA polymerase (EC 2.7.7.7) [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001

C/Accession: AE0064

R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,

Nature 413, 523-527, 2001

Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AE0064

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-789 <RUP>

A/Cross-references: GB:AL590842; PIDN:CA089376.1; PID:G15978613; GSPDB:GN00175

C/Genetics:

A/Gene: YPO0518

C/Superfamily: Escherichia coli DNA-directed DNA polymerase II

C/Keywords: nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 789;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
|||||
DB 546 VIYGDTS 553

RESULT 44

T31321

DNA-directed DNA polymerase (EC 2.7.7.7) B - Cenarchaeum symbiosum

C/Species: Cenarchaeum symbiosum

C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000

C/Accession: T31321

R/Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.

J. Bacteriol. 180, 5003-5009, 1998

Title: Genomic analysis reveals chromosomal variation in natural populations of the ur

A/Reference number: Z20994; MUID:98422450; PMID:9748430

A/Accession: T31321

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-844 <SCH>

A/Cross-references: EMBL:AF083072; NID:G3599393; PID:G3599407; PIDN:AA062712.1

C/Genetics:

A/Gene: pol1

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 844;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 YGPTDSLIF 583
|||||
DB 619 YGPTDSLIF 626

RESULT 45

A56277

DNA-directed DNA polymerase (EC 2.7.7.7) I - Pyrodicticum occultum

C/Species: Pyrodicticum occultum

C/Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 16-Feb-1997

C/Accession: A56277

R/Uemori, T.; Ishino, Y.; Doi, H.; Kato, I.

J. Bacteriol. 177, 2164-2177, 1995

Title: The hyperthermophilic archaeon Pyrodicticum occultum has two alpha-like DNA poly

A/Reference number: A56277; MUID:95238290; PMID:7721707

A/Accession: A56277

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-879 <UEM>

A/Cross-references: GB:D38573

C:Genetics:
A:Gene: polA
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 879;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 576 YGDTDSL 583
DB 654 YGDTDSL 661

RESULT 46

S23019

DNA-directed DNA polymerase (EC 2.7.7.7) - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S23019

Pleasant, F.M.; de Martino, C.; Rossi, M.
C:Title: A DNA polymerase from the archaeon Sulfolobus solfataricus shows sequence simi

C:Reference number: S23019; MUID:92310966; PMID:1614858
A:Accession: S23019

A:Molecule type: DNA
A:Residues: 1-882 <PIS>

A:Cross-references: EMBL:X64466; NID:947564; PIDN:CAA45795.1; PID:947565
A:Experimental source: strain MT 4

C:Genetics:
A:Gene: pols

C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 882;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 576 YGDTDSL 583
DB 653 YGDTDSL 660

RESULT 47

F90201

DNA polymerase I (DNA polymerase B1) (dpol) [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 17-May-2002
C:Accession: F90201

Shen, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
aron, I.; Jettles, A.C.; Kozera, C.D.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
C:Title: The DNA polymerase gene from Sulfolobus solfataricus complete genome.
C:Submitted to Genbank, April 2001

A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139

A:Accession: F90201
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-882 <KUR>

A:Cross-references: GB:AEO06641; NID:913813715; PIDN:AAK40869.1; GSPDB:GN00155
C:Genetics:
A:Gene: dpol

C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 1.0%; Score 8; DB 2; Length 882;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 576 YGDTDSL 583
DB 653 YGDTDSL 660

RESULT 48

T11742

egg sperm receptor - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C:Accession: T11742

R;Foltz, K.R.; Partin, J.S.; Lennarz, W.J.
C:Title: Sea urchin egg receptor for sperm: sequence similarity of binding domain and h

C:Reference number: 217324; MUID:93197888; PMID:8383878
A:Accession: T11742

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

A:Residues: 1-889 <FOL>
A:Cross-references: EMBL:L04969; NID:91580782; PID:91580783

A:Experimental source: immature ovary
C:Superfamily: heat shock protein 91

Query Match 1.0%; Score 8; DB 2; Length 889;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 525 LKVLNAN 532
DB 220 LKVLNAN 227

RESULT 49

A42543

DNA-directed DNA polymerase (EC 2.7.7.7) - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1

C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: A42543

R;Grabherr, R.; Strasser, P.; Van Etten, J.L.
C:Title: The DNA polymerase gene from Chlorella viruses PBCV-1 and NY-2A contains an int

A:Reference number: A42543; MUID:92263776; PMID:1585643
A:Molecule type: DNA

A:Residues: 1-913 <GRA>
A:Cross-references: GB:M86836; NID:9323319; PID:9323320

C:Genetics:
A:Insertion: 682/3

C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA biosynthesis; DNA replication; exonuclease; nucleotidyltran

F;182-199/Region: exonuclease pattern A
F;268-282/Region: exonuclease pattern B
F;388-401/Region: exonuclease pattern C

Query Match 1.0%; Score 8; DB 1; Length 913;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 574 VIYGDTS 581
DB 635 VIYGDTS 642

RESULT 50

B42543

DNA-directed DNA polymerase (EC 2.7.7.7) - Chlorella virus CV-NY-2A
C:Species: Chlorella virus CV-NY-2A

C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 11-Jun-1999
C:Accession: B42543

R;Grabherr, R.; Strasser, P.; Van Etten, J.L.
C:Title: The DNA polymerase gene from Chlorella viruses PBCV-1 and NY-2A contains an int

A:Reference number: A42543; MUID:92263776; PMID:1585643
A:Accession: B42543

A:Molecule type: DNA
A:Residues: 1-913 <GRA>

A:Cross-references: GB:M86837; NID:9323321; PIDN:AAA88827.1; PID:9323322
C:Genetics:

A:introns: 682/3
C:Superfamily: herpesvirus DNA-directed DNA polymerase
C:Keywords: DNA binding; DNA biosynthesis; DNA replication; exonuclease; nucleotidyltran
F:182-199/Region: exonuclease pattern A
F:268-282/Region: exonuclease pattern B
F:388-401/Region: exonuclease pattern C

Query Match

1.0%; Score 8; DB 1; Length 913;
Best Local Similarity 100.0%; Pred. No. 45;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
|||||

Db 635 VIYGDTS 642

RESULT 51

17675

-directed DNA polymerase (EC 2.7.7.7) - Chlorella virus PBCV-1

C:Species: Chlorella virus PBCV-1

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T17675

R:Graves, M.V.; Van Etten, J.L.

Submitted to the EMBL Data Library, May 1999

A:Reference number: Z18806

A/Accession: T17675

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-913 <GRA>

A/Cross-references: EMBL:U42580; NID:94028896; PIDN:AA06553.1

A:Experimental source: specific host Chlorella strain NC64A

C:Genetics:

A:Note: 682/3

A:Introns: 682/3

C:Superfamily: herpesvirus DNA-directed DNA polymerase

C:Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match

1.0%; Score 8; DB 2; Length 913;
Best Local Similarity 100.0%; Pred. No. 45;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
|||||

Db 635 VIYGDTS 642

RESULT 52

2763

Probable DNA-directed DNA polymerase (pfu polymerase) APE0099 - Aeropyrum pernix (strain

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C/Accession: F72763

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takat

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310335; PMID:10382566

A/Accession: F72763

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-959 <KAW>

A/Cross-references: DDBJ:AP000058; NID:95103388; PIDN:BA479008.1; PID:95103487

A:Experimental source: strain K1

C:Genetics:

A:Note: APE0099

C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match

1.0%; Score 8; DB 2; Length 959;
Best Local Similarity 100.0%; Pred. No. 47;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 AFDIEVYS 191
|||||

Db 314 AFDIEVYS 321

RESULT 53

DJBR65

DNA-directed DNA polymerase (EC 2.7.7.7) - human herpesvirus 6 (strain UI102)

C:Species: human herpesvirus 6

A:Note: host Homo sapiens (man)

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 01-Dec-2000

C/Accession: B40898; T09332

R:Teo, I.A.; Griffin, B.E.; Jones, M.D.

J. Virol. 65, 4670-4680, 1991

A:Title: Characterization of the DNA polymerase gene of human herpesvirus 6.

A:Reference number: A40898; MUID:9133007; PMID:1651403

A/Accession: B40898

A:Molecule type: DNA

A:Residues: 1-1012 <TEO>

A/Cross-references: GB:M63804; NID:9325467; PIDN:AAA74631.1; PID:9455196

R:Nicholas, J.; Martin, M.

J. Virol. 68, 597-610, 1994

A:Title: Nucleotide sequence analysis of a 38.5-kilobase-pair region of the genome of hu

A:Reference number: Z16644; MUID:94118404; PMID:8289364

A/Accession: T09332

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 778-1012 <NIC>

A/Cross-references: EMBL:L25528; NID:9451932; PIDN:AAA16745.1; PID:9451963

C:Genetics:

A:Gene: XILF0

C:Superfamily: herpesvirus DNA-directed DNA polymerase

C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match

1.0%; Score 8; DB 1; Length 1012;
Best Local Similarity 100.0%; Pred. No. 50;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
|||||

Db 736 VIYGDTS 743

RESULT 54

T43998

DNA polymerase [imported] - human herpesvirus 6 (strain HST)

C:Species: human herpesvirus 6

A:Variety: strain HST

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C/Accession: T43998

R:Isegawa, Y.; Mukai, T.; Nakano, K.; Kagawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kawa

J. Virol. 73, 8053-8063, 1999

A:Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and

A:Reference number: Z22732; MUID:99412319; PMID:10482554

A/Accession: T43998

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1012 <ISR>

A/Cross-references: EMBL:AB021506; NID:94995977; PIDN:BA478259.1; PID:94996026

A:Experimental source: strain HST; pop. variant B

C:Genetics:

A:Note: U38

C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match

1.0%; Score 8; DB 2; Length 1012;
Best Local Similarity 100.0%; Pred. No. 50;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581
|||||

Db 736 VIYGDTS 743

RESULT 55

T44185

probable DNA-directed DNA polymerase (EC 2.7.7.7) U38 [similarity] - human herpesvirus 6
 C:Species: human herpesvirus 6
 A:Virus: strain 229
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 05-May-2000
 C:Accession: T44185
 R:Domington, G.; Dambach, T.R.; Stamey, F.R.; Dewhurst, S.; Inoue, N.; Pellett, P.E.
 J. Virol. 73, 840-8052, 1999
 A:Title: Human herpesvirus 68 genome sequence: coding content and comparison with human
 A:Reference number: 222734; MUID:99412318; PMID:10482553
 A:Accession: T44185
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1012 <DOM>
 A:Cross-references: EMBL:AF157706; PIDN:AAD49652.1
 A:Experimental source: strain 229; variant B
 C:Genetics:
 A:Note: U38
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1012;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 574 VIYGDTS 581
 |||||
 Db 736 VIYGDTS 743

RESULT 56
 DNA-directed DNA polymerase (EC 2.7.7.7) - Lymantria dispar nuclear polyhedrosis virus
 C:Species: Lymantria dispar nuclear polyhedrosis virus, LdNPV
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 20-Jun-2000
 C:Accession: J01920
 R:Bjornson, R.M.; Glocker, B.; Rohmann, G.F.
 J. Gen. Virol. 73, 3177-3183, 1992
 A:Title: Characterization of the nucleotide sequence of the Lymantria dispar nuclear polyhedrosis virus
 A:Reference number: P00515; MUID:93107852; PMID:1469355
 A:Accession: J01920
 A:Molecule type: DNA
 A:Residues: 1-1013 <BD>
 A:Cross-references: DDBJ:D11476; NID:G1041123; PIDN:BA02036.1; PID:G222230
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1013;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 574 VIYGDTS 581
 |||||
 Db 682 VIYGDTS 689

RESULT 57
 T30431
 DNA-directed DNA polymerase (EC 2.7.7.7) - Lymantria dispar nuclear polyhedrosis virus
 C:Species: Lymantria dispar nuclear polyhedrosis virus, LdNPV
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T30431
 R:Kuzio, V.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohr
 Virol. 73, 17-34, 1999
 A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria d
 A:Reference number: 220836; MUID:99124785; PMID:9887215
 A:Accession: T30431
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1014 <KUZ>
 A:Cross-references: EMBL:AF061810; NID:G3822234; PIDN:AAC70269.1; PID:G3822318
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1014;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 574 VIYGDTS 581
 |||||
 Db 682 VIYGDTS 689

RESULT 58
 T03108
 DNA-directed DNA polymerase (EC 2.7.7.7) - alcelaphine herpesvirus 1
 C:Species: alcelaphine herpesvirus 1
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 18-Jun-1999
 C:Accession: T03108
 R:Enser, A.; Pelanz, R.; Fleckenstein, B.
 J. Virol. 71, 6517-6525, 1997
 A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.
 A:Reference number: Z14840; MUID:97404659; PMID:9261371
 A:Accession: T03108
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1026 <ENS>
 A:Cross-references: EMBL:AF005370; NID:G2337967; PIDN:AAC58060.1; PID:G2337976
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1026;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 576 YGDTSLF 583
 |||||
 Db 770 YGDTSLF 777

RESULT 59
 S19661
 DNA-directed DNA polymerase (EC 2.7.7.7) III large chain - fission yeast (Schizosaccharo
 C:Species: Schizosaccharomyces pombe
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S19661
 R:Pignede, G.; Bouvier, D.; de Recondo, A.M.; Baladaci, G.
 J. Mol. Biol. 222, 209-218, 1991
 A:Title: Characterization of the POI3 gene product from Schizosaccharomyces pombe indica
 A:Reference number: S19661; MUID:92071854; PMID:1960723
 A:Accession: S19661
 A:Molecule type: DNA
 A:Residues: 1-1084 <PIG>
 A:Cross-references: EMBL:X59278; NID:G5010; PIDN:CAA4168.1; PID:G5011
 C:Genetics:
 A:Interons: 77/1
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; nucleotidyltransferase; nucleus

Query Match 1.0%; Score 8; DB 1; Length 1084;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 574 VIYGDTS 581
 |||||
 Db 737 VIYGDTS 744

RESULT 60
 T43266
 DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - fission yeast (Schizosaccharomyc
 N:Alternate names: DNA polymerase delta
 C:Species: Schizosaccharomyces pombe
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T43266
 R:Pignede, G.; Bouvier, D.; de Recondo, A.M.; Baladaci, G.
 J. Mol. Biol. 222, 209-218, 1991

A:Title: Characterization of the POL3 gene product from Schizosaccharomyces pombe indic
 A:Reference number: S19661; MUID:92071954; PMID:11960723
 A:Accession: T43266
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1086 <PIG>
 A:Cross-references: EMBL:L07734; NID:g173383; PIDN:AAA35303.1; PID:g173384
 C:Genetics:
 A:Gene: pol3
 A:Introns: 77/1
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGD TDS 581
 |||||
 737 VIYGD TDS 744

RESULT 61
 T40242
 DNA polymerase delta large chain - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T40242
 R:Bozzym, K.; Beck, A.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999
 A:Reference number: Z21916
 A:Accession: T40242
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1086 <BOR>
 A:Cross-references: EMBL:AL121815; PIDN:CAB58156.1; GSPDB:GN00667; SPDB:SPBC336.04
 C:Experimental source: strain 972h-; cosmid c336
 C:Genetics:
 A:Gene: SPDB:SPBC336.04
 A:Map position: 2
 A:Introns: 77/1
 C:Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 1.0%; Score 8; DB 2; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 574 VIYGD TDS 581
 |||||
 737 VIYGD TDS 744

RESULT 62
 T05731
 DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - soybean
 C:Species: Glycine max (soybean)
 C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 18-Jun-1999
 C:Accession: T05731
 R:Collins, J.T.B.; Cannon, G.C.; Heinrich, S.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z15439
 A:Accession: T05731
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1088 <COL>
 A:Cross-references: EMBL:AF020193; NID:g2895197; PIDN:AAC18443.1; PID:g2895198
 C:Genetics:
 A:Gene: Pol delta
 C:Function:
 A:Description: catalyzes replication of DNA
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1088;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGD TDS 581
 |||||
 Db 714 VIYGD TDS 721

RESULT 63
 S22573
 DNA-directed DNA polymerase (EC 2.7.7.7) delta - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 22-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jun-2000
 C:Accession: S22573; A45639; S17330
 R:Ridley, R.G.; White, J.H.; McAlleese, S.M.; Goman, M.; Alano, P.; de Vries, E.; Kilbey,
 Nucleic Acids Res. 19, 6731-6736, 1991
 A:Title: DNA polymerase delta: gene sequences from Plasmodium falciparum indicate that
 A:Reference number: S22573; MUID:92107655; PMID:1762904
 A:Accession: S22573
 A:Molecule type: DNA
 A:Residues: 1-1094 <RID>
 A:Cross-references: EMBL:X62423; NID:g99940; PIDN:CAA44289.1; PID:g9941
 R:Fox, B.A.; Bizik, D.J.
 Mol. Biochem. Parasitol. 49, 289-296, 1991
 A:Title: The primary structure of Plasmodium falciparum DNA polymerase delta is similar
 A:Reference number: A45639; MUID:92131065; PMID:1775172
 A:Accession: A45639
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1094 <FOX>

A:Cross-references: GB:M64715; NID:g160269; PIDN:AAA29589.1; PID:g160270
 A:Note: Sequence extracted from NCBI backbone (NCBI:77767, NCBI:P.77769)
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: DNA binding; DNA replication; exonuclease; nucleotidyltransferase; nucleus;
 F:1003-1019/Region: zinc finger CCCC motif
 F:1049-1067/Region: zinc finger CCCC motif

Query Match 1.0%; Score 8; DB 2; Length 1094;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGD TDS 581
 |||||
 Db 749 VIYGD TDS 756

RESULT 64
 S40243
 DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S40243; J70670
 R:Cullmann, G.; Hindges, R.; Bertschold, M.W.; Huebscher, U.
 submitted to the EMBL Data Library, March 1993
 A:Reference number: S40243
 A:Accession: S40243
 A:Molecule type: mRNA
 A:Residues: 1-1105 <CUR>
 A:Cross-references: EMBL:Z21848; NID:g438133; PIDN:CAA78995.1; PID:g438134
 R:Cullmann, G.; Hindges, R.; Bertschold, M.W.; Huebscher, U.
 Gene 134, 191-200, 1993
 A:Title: Cloning of a mouse cDNA encoding DNA polymerase delta: refinement of the homolo
 A:Reference number: J70670; MUID:94085777; PMID:8262377
 A:Accession: J70670
 A:Molecule type: DNA
 A:Residues: 1-1111; 'G', '113', 'P', '115-1034', 'Y', '1036-1105' <CU2>
 A:Cross-references: EMBL:Z21848
 A:Note: the sequence translated from Z21848 is inconsistent with that from this sequence
 C:Comment: Three DNA polymerases alpha, delta and epsilon chains are essential. This de
 C:Genetics:
 A:Gene: poldelta
 C:Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA replication; heterodimer; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 1; Length 1105;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 621 KRRYAGLL 628
Db 804 KRRYAGLL 811

RESULT 65

A39299
DNA-directed DNA polymerase (EC 2.7.7.7) delta catalytic chain - bovine

N/Alternate names: DNA-directed DNA polymerase III 125K chain

N/Contains: DNA-directed DNA polymerase delta, 116K chain

C/Species: Bos primigenius taurus (cattle)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: A39299

R/Chang, J.; Chung, D.W.; Tan, C.K.; Downey, K.M.; Davie, E.W.; So, A.G.

Chemistry 30, 11742-11750, 1991

Title: Primary structure of the catalytic subunit of calf thymus DNA polymerase delta:

Reference number: A39299; PMID:92089082; PMID:1721537

A/Accession: A39299

A/Molecule type: mRNA

A/Residues: 1-1106 <ZHA>

A/Cross-references: GB:M80395; NID:G162973; PIDN:AAA30493.1; PID:G162974

A/Experimental source: thymus

A/Note: 131-Glu and 263-Ser were also found by mRNA sequencing

A/Note: Part of this sequence was confirmed by protein sequencing

A/Note: The amino end of the mature protein was blocked

C/Comment: A 116K polypeptide was shown to be a degradation product of the mature protein

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; nucleus; zinc finger

Query Match

Best Local Similarity 1.0%; Score 8; DB 1; Length 1106;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 621 KRRYAGLL 628
Db 805 KRRYAGLL 812

RESULT 66

A41618

DNA-directed DNA polymerase (EC 2.7.7.7) delta catalytic chain - human

C/Species: Homo sapiens (man)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: A41618; S35455

R/Chung, D.W.; Zhang, J.; Tan, C.K.; Davie, E.W.; So, A.G.; Downey, K.M.

Proc. Natl. Acad. Sci. U.S.A. 88, 11197-11201, 1991

Title: Primary structure of the catalytic subunit of human DNA polymerase delta and cH

Reference number: A41618; PMID:92107916; PMID:1722322

A/Accession: A41618

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1107 <CHU>

A/Cross-references: GB:M80397; NID:G181619; PIDN:AAA58439.1; PID:G181620

R/Yang, C.L.; Chang, L.S.; Zhang, P.; Hao, H.; Zhu, L.; Toomey, N.L.; Lee, M.Y.W.T.

Nucleic Acids Res. 20, 735-745, 1992

Title: Molecular cloning of the cDNA for the catalytic subunit of human DNA polymerase

Reference number: S35455; PMID:92178967; PMID:1542570

A/Accession: S35455

A/Molecule type: mRNA

A/Residues: 1-129 'R', 31-118 'H', 120-172 'N', 174-471 'V', 473-775 'G', 777-1107 <YAN>

A/Cross-references: EMBL:M81735; NID:G181621; PIDN:AAA35768.1; PID:G181622

C/Genetics

A/Genes: GDB:PODI, POD

A/Cross-references: GDB:129089; OMIM:174761

A/Map position: 19q13.3-19q13.3

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; nucleus; zinc finger

Query Match 1.0%; Score 8; DB 1; Length 1107;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 621 KRRYAGLL 628
Db 806 KRRYAGLL 813

RESULT 67

DJBE3

DNA-directed DNA polymerase (EC 2.7.7.7) - equine herpesvirus 1 (strain Ab4p)

C/Species: equine herpesvirus 1

A/Note: host Equus caballus (domestic horse)

C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999

C/Accession: D36798

R/Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

submitted to GenBank, March 1992

A/Description: The DNA sequence of equine herpesvirus-1.

A/Reference number: A36805

A/Accession: D36798

A/Molecule type: DNA

A/Residues: 1-1120 <TEL>

A/Cross-references: GB:M86664; NID:G330791; PIDN:AA02465.1; PID:G330822

R/Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

Virology 189, 104-116, 1992

Title: The DNA sequence of equine herpesvirus-1.

A/Reference number: A41811; PMID:9229566; PMID:1318606

A/Contents: annotation; possible protein-coding frames

A/Note: neither amino acid nor nucleotide sequence is given

C/Genetics

A/Genes: 30

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match

Best Local Similarity 1.0%; Score 8; DB 1; Length 1220;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 574 VIYGDFTS 581
Db* 883 VIYGDFTS 890

RESULT 68

T42573

DNA-directed DNA polymerase (EC 2.7.7.7) - equine herpesvirus 4 (strain NS80567)

C/Species: equine herpesvirus 4

A/Variety: strain NS80567

C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C/Accession: T42573

R/Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.

J. Gen. Virol. 79, 1197-1203, 1998

Title: The DNA sequence of equine herpesvirus-4.

A/Reference number: 222173; PMID:98264497; PMID:9603335

A/Accession: T42573

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1120 <TEL>

A/Cross-references: EMBL:AF030027; NID:G260550; PIDN:AA059546.1; PID:G2605974

A/Experimental source: strain NS80567

C/Genetics

A/Genes: 30

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 1220;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 574 VIYGDFTS 581
|||||||

Db 884 VIYGDTS 891

RESULT 69

DUBBC1

DNA-directed DNA polymerase (EC 2.7.7.7) - human cytomegalovirus (strain AD169)

N/Alternate names: HFLP2 protein; UL54 protein

C/Species: human cytomegalovirus, human herpesvirus 5

A/Note: host Homo sapiens (man)

C/Date: 31-Mar-1988 #sequence_revision 31-Dec-1990 #text_change 11-Jun-1999

A/Accession: S09817; A25983

R/Chen, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Hornsnel, T.; M.; Barrell, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A/Reference number: S09749; MUID:90269039; PMID:2161319

A/Accession: S09817

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-1242 <CHE>

A/Cross-references: EMBL:X17403; NID:G59591; PIDN:CAA35413.1; PID:G1780832

A/Note: possible protein-coding frames are given

R/Note: The DNA sequence was submitted to EMBL, December 1989, in computer-readable form

R/Kouzarides, T.; Bankier, A.T.; Satchwell, S.C.; Weston, K.; Tomlinson, P.; Barrell, B.

J. Virol. 61, 125-133, 1987

A/Title: Sequence and transcription analysis of the human cytomegalovirus DNA polymerase

A/Reference number: A25983; MUID:87061230; PMID:3023650

A/Accession: A25983

A/Molecule type: DNA

A/Residues: 1-1242 <KOU>

A/Cross-references: GB:M14709; NID:G330640; PIDN:AAA45988.1; PID:G330642

C/Genetics:

A/Map position: unique long region (U-L)

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 1.0%; Score 8; DB 1; Length 1242;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581

Db 906 VIYGDTS 913

RESULT 70

J278

DNA fiber protein gp37 - phage K3

C/Species: phage K3

C/Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 20-Sep-1999

C/Accession: S07278; I41090; I41091

R/Riede, I.; Drexler, K.; Eschbach, M.L.; Henning, U.

J. Mol. Biol. 191, 255-266, 1996

A/Title: DNA sequence of the tail fiber genes 37, encoding the receptor recognizing part

A/Reference number: S07278; MUID:87112716; PMID:3806672

A/Accession: S07278

A/Molecule type: DNA

A/Residues: 1-1243 <RIE>

A/Cross-references: EMBL:X04747; NID:G15110; PIDN:CAA28445.1; PID:G15111

R/Riede, I.; Eschbach, M.L.; Henning, U.

Mol. Gen. Genet. 195, 144-152, 1984

A/Title: DNA sequence heterogeneity in the genes of T-even type Escherichia coli phages

A/Reference number: I41090; MUID:85035820; PMID:6052843

A/Accession: I41090

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 158-211 <RES>

A/Cross-references: EMBL:X00613; NID:G41844; PIDN:CAA25250.1; PID:G929577

A/Accession: I41091

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 727-792 <RE2>

A/Cross-references: EMBL:X00615; NID:G41845; PIDN:CAA25251.1; PID:G929578

C/Genetics:
A/Genes: 37
C/Superfamily: phage T4 tail fiber protein gp37
C/Keywords: dimer; tail fiber

Query Match 1.0%; Score 8; DB 2; Length 1243;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 VLPEAVRE 102

Db 1196 VLPEAVRE 1203

RESULT 71

S20052

DNA-directed DNA polymerase (EC 2.7.7.7) alpha catalytic chain - Trypanosoma brucei brucei

C/Species: Trypanosoma brucei brucei

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: S20052; S15729

R/Deegater, P.A.J.; Strating, M.; Murphy, N.B.; Kooy, R.F.; van der Vliet, P.C.; Overdu

Nucleic Acids Res. 19, 6441-6447, 1991

A/Title: The Trypanosoma brucei DNA polymerase alpha core subunit gene is developmentally

A/Reference number: S20051; MUID:92093600; PMID:1754381

A/Accession: S20052

A/Molecule type: DNA

A/Residues: 1-1339 <LEB>

A/Cross-references: EMBL:X60951; NID:G10499; PIDN:CAA43287.1; PID:G10501

A/Note: 828-Arg was also found

C/Genetics:

A/Genes: pola

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; nucleus; zinc finger

Query Match 1.0%; Score 8; DB 1; Length 1339;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581

Db 894 VIYGDTS 901

RESULT 72

JC5508

DNA-directed DNA polymerase (EC 2.7.7.7) alpha catalytic chain - Leishmania donovani

C/Species: Leishmania donovani

C/Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 18-Jun-1999

C/Accession: JC5508

R/Luton, K.; Johnson, A.M.

Biochem. Biophys. Res. Commun. 234, 95-100, 1997

A/Title: Cloning and sequence analysis of the DNA polymerase alpha gene of Leishmania dc

A/Reference number: JC5508; MUID:97312527; PMID:9168968

A/Accession: JC5508

A/Molecule type: DNA

A/Residues: 1-1339 <LUT>

A/Cross-references: GB:U78172; NID:G2138287; PIDN:AAQ47538.1; PID:G2138288

C/Superfamily: herpesvirus DNA-directed DNA polymerase

C/Keywords: nucleotidyltransferase

Query Match 1.0%; Score 8; DB 2; Length 1339;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDTS 581

Db 896 VIYGDTS 903

RESULT 73

S09579

tail fiber protein gp37 - phage T2

C/Species: phage T2

C>Date: 10-Mar-1994 #sequence_revlelon 10-Mar-1994 #text_change 20-Sep-1999
 C/Accession: S09579
 R:Riede, I.; Drexler, K.; Eschbach, M.L.; Henning, U.
 J. Mol. Biol. 191, 255-266, 1986
 A>Title: DNA sequence of the tail fiber genes 37, encoding the receptor recognizing part
 A/Reference number: S07278; MUID:87112716; PMID:360672
 A/Accession: S09579
 A/Molecule type: DNA
 A/Residues: 11341 <RES>
 A/Cross-references: EMBL:X04442; NID:915195; PIDN:CAA28038.1; PID:915196
 C/Genetics:
 A/Genes: 37
 C/Superfamily: phage T4 tail fiber protein gp37
 C/Keywords: dimer; tail fiber

Query Match 1.0%; Score 8; DB 2; Length 1341;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 VLPEAVRE 102
 |||||
 1294 VLPEAVRE 1301

RESULT 74

DJ2PA
 DNA-directed DNA polymerase (EC 2.7.7.7) alpha - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 10-Dec-1999
 C/Accession: S15993; F38752; S21818
 R:Dimaquer, V.; Tillet, J.; de Rendo, A.M.; Baldacci, G.
 Mol. Gen. Genet. 226, 182-189, 1991
 A>Title: The POL1 gene from the fission yeast, Schizosaccharomyces pombe, shows conserve
 A/Reference number: S15993; MUID:91238692; PMID:2034212
 A/Accession: S15993
 A/Molecule type: DNA
 A/Residues: 1-1405 <MOL>
 A/Cross-references: EMBL:X58299; NID:95008; PIDN:CAA1232.1; PID:95009
 R:Oliver, K.; Harris, D.; Wood, V.; Bartell, B.G.; Rajandream, M.A.
 Submitted to the EMBL Data Library, September 1996
 A/Reference number: Z21809
 A/Accession: T38752
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1405 <OLI>
 A/Cross-references: EMBL:Z99296; PIDN:CAB16598.1; GSPDB:GNO0066; SPDB:SPAC3H5.06c
 A/Experimental source: strain 972b-; cosmid c3H5
 C/Genetics: <MOL1>
 Gene: POL1
 Map position: II
 Genetics: <OLI2>
 A/Genes: SPAC3H5.06c
 A/Map position: 1
 A/Intons: 101/1
 C/Superfamily: herpesvirus DNA-directed DNA polymerase
 C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; zinc finger
 F:116-1340/Region: zinc finger CCCC motif

Query Match 1.0%; Score 8; DB 1; Length 1405;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDITS 581
 |||||
 Db 978 VIYGDITS 985

RESULT 75

DJH0AC
 DNA-directed DNA polymerase (EC 2.7.7.7) alpha catalytic chain - human
 C/Species: Homo sapiens (man)
 C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 11-Jun-1999
 C/Accession: S00257; A30440; T57513; S12665

R:Mong, S.W.; Wahl, A.F.; Yuan, P.M.; Arai, N.; Pearson, B.E.; Arai, K.I.; Korn, D.; Hu
 EMBD 37, 37-47, 1988
 A>Title: Human DNA polymerase alpha gene expression is cell proliferation dependent and
 A/Reference number: S00257; MUID:88196090; PMID:3359994
 A/Accession: S00257
 A/Molecule type: mRNA
 A/Residues: 1-1462 <MON>
 A/Cross-references: EMBL:X06745; NID:935567; PIDN:CAA29920.1; PID:935568
 A/Accession: A30440
 A/Molecule type: protein
 A/Residues: 438-449; 495-502; 'G', 838-848; 1090-1105; 1201-1216; 1397-1407; 1444-1453 <MON2>
 R:Pearson, B.E.; Nashner, H.P.; Wang, T.S.
 Mol. Cell. Biol. 11, 2081-2095, 1991
 A>Title: Human DNA polymerase alpha gene: sequences controlling expression in cycling an
 A/Reference number: 157513; MUID:91172197; PMID:2005899
 A/Accession: 157513
 A/Status: translation not shown; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-8 <RES>
 A/Cross-references: GB:M64481; NID:9181617; PIDN:AAA52318.1; PID:9181618
 R:Hei, K.L.; Copeland, W.C.; Wang, T.S.F.
 Nucleic Acids Res. 18, 6231-6237, 1990
 A>Title: Human DNA polymerase alpha catalytic polypeptide binds ConA and RCA and contrai
 A/Reference number: S12665; MUID:91057099; PMID:2243771
 A/Accession: S12665
 A/Molecule type: protein
 A/Residues: 19-37; 'C', 1406-1425, 'C' <HSI>
 C/Genetics:
 A/Genes: GDB:POLA
 A/Cross-references: GDB:120304; OMTM:312040
 A/Map position: XP22.3-Xp21.1
 C/Superfamily: herpesvirus DNA-directed DNA polymerase
 C/Keywords: DNA binding; DNA replication; nucleotidyltransferase; zinc finger
 F:554-691/Region: zinc finger CHCC motif
 F:1249-1374/Region: zinc fingers

Query Match 1.0%; Score 8; DB 1; Length 1462;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 VIYGDITS 581
 |||||
 Db 998 VIYGDITS 1005

Search completed: April 23, 2003, 07:50:01
 Job time : 48 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 07:28:12 ; Search time 29 Seconds
(without alignments)
814.711 Million cell updates/sec

Title: US-10-034-849-2

Perfect score: 4188

Sequence: 1 MTEVFTVLDSSEYVGVKPEP.....LKASATGQKTLFPDLAKKSK 803

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5a.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4188	100.0	803	2	US-08-907-166-4
2	3028.5	72.3	803	1	US-08-062-368-4
3	2963.5	70.8	803	1	US-08-062-368-2
4	1563	37.3	788	2	US-08-907-166-6
5	1237.5	29.5	774	2	US-08-902-632-2
6	1237.5	29.5	776	2	US-08-688-649-37
7	1236.5	29.5	775	1	US-07-966-278-1
8	1236.5	29.5	775	1	US-08-424-921-1
9	1236.5	29.5	775	2	US-08-556-355A-1
10	1236.5	29.5	775	2	US-07-803-627A-1
11	1234.5	29.5	774	3	US-09-073-354-1
12	1234.5	29.5	774	3	US-08-656-005A-1
13	1234.5	29.5	774	4	US-09-073-259-1
14	1234.5	29.5	774	4	US-09-363-095-1
15	1234.5	29.5	774	4	US-09-418-027-1
16	1146.5	27.4	778	2	US-08-906-925-4
17	1104	26.4	779	1	US-08-375-134-12
18	1104	26.4	779	5	PCT-US95-15263-12
19	760.5	18.2	1107	1	US-08-366-577-2
20	760.5	18.2	1107	5	PCT-US96-00005-2
21	739.5	17.7	1019	1	US-08-271-364A-7
22	739.5	17.7	1019	2	US-08-223-715B-26
23	635	15.2	1022	1	US-08-271-364A-8
24	635	15.2	1022	2	US-08-223-715B-27
25	579.5	13.8	1462	3	US-07-792-600-31
26	579.5	13.8	1462	3	US-09-157-021-31
27	579.5	13.8	1462	3	US-09-156-842-31

28	542.5	13.0	877	2	US-08-907-166-8	Sequence 8, Appl
29	524.5	12.5	1015	2	US-08-680-326-32	Sequence 32, Appl
30	515	12.3	1009	2	US-08-680-326-31	Sequence 31, Appl
31	506.5	12.1	762	2	US-08-907-166-10	Sequence 10, Appl
32	502.5	12.0	837	2	US-08-680-326-117	Sequence 117, Appl
33	495	11.8	1008	2	US-08-680-326-30	Sequence 30, Appl
34	458.5	10.9	396	1	US-08-229-284A-2	Sequence 2, Appl
35	452	10.8	1240	2	US-08-680-326-37	Sequence 37, Appl
36	450.5	10.8	1220	2	US-08-680-326-38	Sequence 38, Appl
37	450	10.7	1194	2	US-08-680-326-35	Sequence 35, Appl
38	444	10.6	1235	2	US-08-680-326-36	Sequence 36, Appl
39	440.5	10.5	783	1	US-08-101-593-6	Sequence 6, Appl
40	440.5	10.5	783	1	US-08-465-995A-6	Sequence 6, Appl
41	440.5	10.5	783	2	US-08-465-994C-6	Sequence 6, Appl
42	424	10.1	1012	2	US-08-680-326-34	Sequence 34, Appl
43	419	10.0	1094	2	US-08-680-326-40	Sequence 40, Appl
44	394	9.4	1097	2	US-08-680-326-39	Sequence 39, Appl
45	368	8.8	1242	2	US-08-680-326-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1									
US-08-907-166-4									
Sequence 4, Application US/08907166									
Patent No. 5948666									
GENERAL INFORMATION:									
APPLICANT: Callen, Walter									
* APPLICANT: Mather, Eric									
TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES									
FILE REFERENCE: 09010/027001									
CURRENT APPLICATION NUMBER: US/08/907,166									
NUMBER OF SEQ ID NOS: 12									
CURRENT FILING DATE: 1997-08-06									
SOFTWARE: FastSeq for Windows Version 3.0									
SEQ ID NO 4									
LENGTH: 803									
TYPE: PRT									
ORGANISM: Pyrobolus fumarius									
US-08-907-166-4									
Query Match									
Best Local Similarity 100.0%; Score 4188; DB 2; Length 803;									
Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MTEVFTVLDSSEYVGVKPEPQVIMGIAENGGERVYLIDRSFRFYFALLAPGADPKQVAQ	60						
DB	1	MTEVFTVLDSSEYVGVKPEQVIMGIAENGGERVYLIDRSFRFYFALLAPGADPKQVAQ	60						
QY	61	RIRALSPKSPITIGVEDDKRKYFGRRPVLRITVLEAVREYRELKQNVGVEDVLEAD	120						
DB	61	RIRALSPKSPITIGVEDDKRKYFGRRPVLRITVLEAVREYRELKQNVGVEDVLEAD	120						
QY	121	IRFMRRLIDHDLPFPMVYVEAPLENKGFVVDKYLKSPREPLYGALAPTKLPDL	180						
DB	121	IRFMRRLIDHDLPFPMVYVEAPLENKGFVVDKYLKSPREPLYGALAPTKLPDL	180						
QY	181	RILAFDLEVYSKQSPREPRDPVILAVKTDGDEVFLIAGKQDPRKPIEFPEYVGRYD	240						
DB	181	RILAFDLEVYSKQSPREPRDPVILAVKTDGDEVFLIAGKQDPRKPIEFPEYVGRYD	240						
QY	241	PDIIVGNNHFPDMPYLRRARILIGIKLDVTRVGAEPPTSVMGHVSPGRNLVDLYDA	300						
DB	241	PDIIVGNNHFPDMPYLRRARILIGIKLDVTRVGAEPPTSVMGHVSPGRNLVDLYDA	300						
QY	301	EMPEIKIKSLAEVAVLVGMKKSERYITIMWEIPDVWDPKRPIILQYARDVATYGS	360						
DB	301	EMPEIKIKSLAEVAVLVGMKKSERYITIMWEIPDVWDPKRPIILQYARDVATYGS	360						
QY	361	LAEKILPFAQLSVTGLPDDQVAMSVGRLEEMYLIRAFKMKELVKNVPEEETRYG	420						
DB	361	LAEKILPFAQLSVTGLPDDQVAMSVGRLEEMYLIRAFKMKELVKNVPEEETRYG	420						

Wed Apr 23 08:09:10 2003

us-10-034-849-2.ra1

Page 2

Qy 421 AIVLEPLRGVHENIATLDSSMYPMIKYXNVGPDITVPRGKCGCGWEPVYKRRFR 480
Db 421 AIVLEPLRGVHENIATLDSSMYPMIKYXNVGPDITVPRGKCGCGWEPVYKRRFR 480
Qy 481 RCPGPFKTVLELLELRKRVAAEMKKTPPSPPEYRLDERQALKVLANASYGMGSG 540
Db 481 RCPGPFKTVLELLELRKRVAAEMKKTPPSPPEYRLDERQALKVLANASYGMGSG 540
Qy 541 ARMYCBKAAVATWGRHLIRTAIIRAKGLKVIYGDTSLEFVYDPEKVENFKI1KE 600
Db 541 ARMYCBKAAVATWGRHLIRTAIIRAKGLKVIYGDTSLEFVYDPEKVENFKI1KE 600
Qy 601 ELGFEIKLEKVKYRLEFTEAKKRYAGLLEDRIDIVGFVAVRGDMCELAKEVQVVEIV 660
Db 601 ELGFEIKLEKVKYRLEFTEAKKRYAGLLEDRIDIVGFVAVRGDMCELAKEVQVVEIV 660
Qy 661 LKTSVNVKAVEYRKIVKLEEGKVPLEKVIWKTLSKRLSEYTTAPHVAAKRMISAG 720
Db 661 LKTSVNVKAVEYRKIVKLEEGKVPLEKVIWKTLSKRLSEYTTAPHVAAKRMISAG 720
Qy 721 YRVSPPGKIGYIVKGGGRISORAMPYPMVDPDIDVTYVVDHOIIPAAIRILGRTIT 780
Db 721 YRVSPPGKIGYIVKGGGRISORAMPYPMVDPDIDVTYVVDHOIIPAAIRILGRTIT 780
Qy 781 EKLKASATGOKTLPDLFLAKSK 803
Db 781 EKLKASATGOKTLPDLFLAKSK 803

RESULT 2

US-08-062-368-4
Sequence 4, Application US/08062368
Patent No. 5491086
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
TITLE OF INVENTION: Purified Thermostable Nucleic Acid
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/062.368
FILING DATE: 19930514
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sias Ph.D., Stacey R.
REGISTRATION NUMBER: 32,630
REFERENCE/DOCKET NUMBER: 8584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2863
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 803 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-062-368-4

Query Match

72.3%, Score 3028.5, DB 1, Length 803;

Best Local Similarity 70.8%; Pred. No. 1e-270;
Matches 570; Conservative 106; Mismatches 122; Indels 7; Gaps 5;

Qy 1 MTEVU-FTVLDSSYEVVGEKPOVITWGLAENGERRVTLIDNSFRFPYALALGAD--KQ 57
Db 1 MTEVU-FTVLDSSYEVVGEKPOVITWGLAENGERRVTLIDNSFRFPYALALGAD--KQ 57
Qy 58 VAQIRIALSRKSPDIIGVEDDKRYGPRVRLIRTVLPEAVAEYELKVNVDVEDVL 117
Db 61 IASIRRLSVKSPDIIDAKPLDRYGRPRKVKTTTMIPESVYHVEAAVKLEGVESDL 120
Qy 118 EADIFPARKYLIDDLPEFTWVREAEPLNNMGFRVVKVLYVSRREPLVYGAALPTQL 177
Db 121 EADIFPARKYLIDRLPFTVYRIPEVDAGRPRGFRVRYKVAADPEPL--ADITRIDL 178
Qy 178 PDRLIAPDIEVYSKQSPRPERDPVIVAVKTDGDDEVLPIAGKDRKCPREFEYVK 237
Db 179 PEMRLVADIEVYSRGRSPNADPVITVLSLDSGKRLLEAGCHDRVLRLEFEYVR 238
Qy 238 RYDPLIYGNMHPMPYLLRPAIRILGKLDYPRVGAEPFYSVGHVSPGRVLDVLY 297
Db 239 ADPDLIYGNMHPMPYLLRPAIRILGKLDYPRVGAEPFYSVGHVSPGRVLDVLY 298
Qy 298 DVAEMPEIKLSLEVAEYLSYMKSSERVITINMEIDYVDDPKKRPILLQYARDVRA 357
Db 299 DVAEMPEIKLSLEVAEYLSYMKSSERVITINMEIDYVDDPKKRPILLQYARDVRA 358
Qy 358 TYGLAEKILPPAIOISYVTGLPLDQVANSVGFLEWYLIRAAFKMELVNNRVERPEET 417
Db 359 TYGLAEKMLPPAIOISYVTGLPLDQVANSVGFLEWYLIRAAFKMELVNNRVERGES 418
Qy 418 YRGAIVLEPLRGVHENIATLDSSMYPMIKYXNVGPDITVPRGKCGCGWEPVYKRRFR 476
Db 419 YRGAIVLEPLRGVHENIATLDSSMYPMIKYXNVGPDITVPRGKCGCGWEPVYKRRFR 478
Qy* 477 HFFRRCPPGPFKTVLELLELRKRVAAEMKKTPPSPPEYRLDERQALKVLANASYGM 536
Db 479 HFFRRCPPGPFKTVLELLELRKRVAAEMKKTPPSPPEYRLDERQALKVLANASYGM 538
Qy 537 GMSGARWYCBKAAVATWGRHLIRTAIIRAKGLKVIYGDTSLEFVYDPEKVENFKI 596
Db 539 GMSGARWYCBKAAVATWGRHLIRTAIIRAKGLKVIYGDTSLEFVYDPEKVENFKI 598
Qy 597 IIKELGFEIKLEKVKYRLEFTEAKKRYAGLLEDRIDIVGFVAVRGDMCELAKEVQV 656
Db 599 IIKELGFEIKLEKVKYRLEFTEAKKRYAGLLEDRIDIVGFVAVRGDMCELAKEVQV 658
Qy 657 VEIVLKTSEVNVKAVEYRKIVKLEEGKVPLEKVIWKTLSKRLSEYTTAPHVAAKRM 716
Db 659 VEIVLKTSEVNVKAVEYRKIVKLEEGKVPLEKVIWKTLSKRLSEYTTAPHVAAKRM 718
Qy 717 LSAQYRVSPPGKIGYIVKGGGRISORAMPYPMVDPDIDVTYVVDHOIIPAAIRILG 776
Db 719 LSAQYRVSPPGKIGYIVKGGGRISORAMPYPMVDPDIDVTYVVDHOIIPAAIRILG 777
Qy 777 FGITEKLLKASATGOKTLPDLFLAKSK 801
Db 778 FGITEKLLKASATGOKTLPDLFLAKSK 802

RESULT 3

US-08-062-368-2
Sequence 2, Application US/08062368
Patent No. 5491086
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
TITLE OF INVENTION: Purified Thermostable Nucleic Acid
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street

CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/062.368
FILING DATE: 19930514
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sias Ph.D., Stacey R.
REGISTRATION NUMBER: 32,630
REFERENCE/DOCKET NUMBER: 8584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2863
TELEFAX: (510) 814-2977
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 803 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-062-368-2

Query Match 70.8%; Score 2963.5; DB 1; Length 803;
Best Local Similarity 68.4%; Pred. No. 1e-264;
Matches 551; Conservative 118; Mismatches 129; Indels 7; Gaps 5;

QY 1 MTEVVFVTLVDSSEYVNGKEPQVITWIGIAENGERVVLIDRSFRPYFALLAPGADPK--Q 57
DB 1 MPEAIEFVLDSSEYVNGKEPQVITWIGIAENGERVVLIDRSFRPYFALLAPGADPK--Q 60
QY 58 VAQIRALSRPKSPIIGVEDDKRYFGRRVLRIRTVLPEAVREYRLVKNVGDVEVL 117
DB 61 VVAIRRLSMKSPILIEKVVSKYFGRRVLRIRTVLPEAVREYRLVKNVGDVEVL 120
QY 118 EADIRFAMRYLIDHLPFTWYRVAEPLNKGFRVDKVLVKSREPLEYGEALAPTKL 177
DB 121 EADIRFAMRYLIDHLPFTWYRVAEPLNKGFRVDKVLVKSREPLEYGEALAPTKL 178
QY 178 PDLRIALFIDIEVYKSGSPRPDPVIVIAVKTDDGDEVLFIABGKODRPIREFVEYK 237
DB 179 PEMRVLAFLDIEVYKSGSPRPDPVIVIAVKTDDGDEVLFIABGKODRPIREFVEYK 238
QY 238 RYDDIIVGNNHFPDMPYLLRARIIGIKLDVTRVGAEPPTSVAHGVSFGRNLVDLY 297
DB 239 SFDDIIVGNNHFPDMPYLLRARIIGIKLDVTRVGAEPPTSVAHGVSFGRNLVDLY 298
QY 298 DYAEEMPEIKISLEEVAYELGVKSKSERVIINWMEIPDYMDPKRPLLLQYARADVRA 357
DB 299 NYVEEMPEIKISLEEVAYELGVKSKSERVIINWMEIPDYMDPKRPLLLQYARADVRA 358
QY 358 TYGLAEKILPFAIOLSVYTGPLDQVGMASVGRLEWYLIRAFKMKELVNNRVERPEET 417
DB 359 TYGLAEKILPFAIOLSVYTGPLDQVGMASVGRLEWYLIRAFKMKELVNNRVERPEET 418
QY 418 YRGALVLEPLRGVHENIAVLDFSSMYPNIMIKYVNGPDTLVVRPEKGE--CGCEAPAEYK 476
DB 419 YRGALVLEPLRGVHENIAVLDFSSMYPNIMIKYVNGPDTLVVRPEKGE--CGCEAPAEYK 478
QY 477 HRFRCRPGPFKTYLERLLELRKVRAMKCYPPDSPEYRLDROKALKYLANASYGYM 536
DB 479 HRFRCRPGPFKTYLERLLELRKVRAMKCYPPDSPEYRLDROKALKYLANASYGYM 538
QY 537 GWSGARVYCRCAKAVTAMGHLRTAINARKGLKTYIGDTSLEFVYDPEKVENFIK 596
DB 539 GWSGARVYCRCAKAVTAMGHLRTAINARKGLKTYIGDTSLEFVYDPEKVENFIK 598

QY 597 IIKBELGFEIKLEKYYKRLFEATEAKKRYAGLLEDGRIDIVGFPAVGRMDCELAKEYOTKY 656
DB 599 FVEKQJGFEIKIDKVKRYRFFTEAKRYVGLLEDGRMDIVGFPAVGRMDCELAKEYOTKY 658
QY 657 VEIVLKTSEVNAKAVEYRKIVKLEBGKVPDIEKVIWIKTLSKRLSEYTTAPVNAKRM 716
DB 659 ABIIKKTGIDINRAISYIREVVRRLRGSKIPITLVIMKLTIKRIEYBHAHPVTAARM 718
QY 717 LSGCYRVSPEDKIGYIVKGGGRISQAMPYEVKQPSQIDVYVYDHOIIPALRILGY 776
DB 719 KEAGYDAPGDKIGYIVKGGGRISQAMPYEVKQPSQIDVYVYDHOIIPALRILGY 777
QY 777 FGITEKKTKASAGOKTLPEDFLAK 801
DB 778 FGITEKKTKASAGOKTLPEDFLAK 802

RESULT 4

US-08-907-166-6
Sequence 6, Application US/08907166
Patent No. 594866
GENERAL INFORMATION:
APPLICANT: Callen, Walter
APPLICANT: Mather, Eric
TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
FILE REFERENCE: 09010/027001
CURRENT APPLICATION NUMBER: US/08/907,166
NUMBER OF SEQ ID NOS: 12
CURRENT FILING DATE: 1997-08-06
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 788
TYPE: PRT
ORGANISM: Archaeoglobus lithotrophicus
US-08-907-166-6

Query Match 37.3%; Score 1563; DB 2; Length 788;
Best Local Similarity 41.3%; Pred. No. 2e-135;
Matches 338; Conservative 154; Mismatches 273; Indels 54; Gaps 18;

QY 1 MTEVVFVTLVDSSEYVNGKEPQVITWIGIAENGERVVLIDRSFRPYFALLAPGADPK--Q 60
DB 1 MTKVGMGLDADITENDRAVIRLMCKDEBG--IFIAVDHSFOFYALKEGITAADIYK 59
QY 61 -RIRALSREPKSPIIGVEDDKRYFGRRVLRIRTVLPEAVREYRLVKNVGDVEVL 119
DB 60 IKVQTKKEVITP-LKVEETAKNIGREVEVFKIYARHPQVPLREVSG--YLEIRBA 115
QY 120 DIRFAMRYLIDHLPFTWYRVAEPLNKGFRVDKVLVKSREPLEYGEALAPTKL 177
DB 116 DIRFAMRYLIDHLPFTWYRVAEPLNKGFRVDKVLVKSREPLEYGEALAPTKL 177
QY 178 PDLRIALFIDIEVYKSGSPRPDPVIVIAVKTDDGDEVLFIABGKODRPIREFVEYK 237
DB 164 PELKMAFDEMLSEVGMPEPKDPIIVISIKSGYEELI--NGDNERELTRFVKIIR 220
QY 238 RYDDIIVGNNHFPDMPYLLRARIIGIKLDVTRVGAEPPTSVAHGVSFGRNLVDLY 297
DB 221 DIEDPIIVGNNHFPDMPYLLRARIIGIKLDVTRVGAEPPTSVAHGVSFGRNLVDLY 298
QY 298 DYAEEMPEIKISLEEVAYELGVKSKSERVIINWMEIPDYMDPKRPLLLQYARADVRA 357
DB 279 DYAEEMPEIKISLEEVAYELGVKSKSERVIINWMEIPDYMDPKRPLLLQYARADVRA 357
QY 358 TYGLAEKILPFAIOLSVYTGPLDQVGMASVGRLEWYLIRAFKMKELVNNRVERPEET 417
DB 336 TYGLAEKILPFAIOLSVYTGPLDQVGMASVGRLEWYLIRAFKMKELVNNRVERPEET 417
QY 418 YRGALVLEPLRGVHENIAVLDFSSMYPNIMIKYVNGPDTLVVRPEKGE--CGCEAPAEYK 477
DB 395 YEGAFVLEPARGLHENIYCLDFASMPYSIMISYNISPDTLVI--GKCDCCNV--ADEVGH 450
QY 478 RFRRCRPGPFKTYLERLLELRKVRAMKCYPPDSPEYRLDROKALKYLANASYGYM 537

CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/688,649
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 220/281
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-688-649-37

Query Match 29.5%; Score 1237.5; DB 2; Length 776;
Best Local Similarity 36.1%; Pred. No. 2.2e-105;
Matches 301; Conservative 154; Mismatches 283; Indels 95; Gaps 23;

QY 8 VLDSSVEVNGKEPQVITWGAENGERVVLIDRSFEPYFALLAGADPKQV---AQRIR 63
DB 2 ILDDVYITEBCKPVIRLFK-KENGRKFIHRTFRPIYALLRDSKIEBKKTGERHG 60
QY 64 ALSRPKPIIGVEDDKKRYGRRPRVLRIRTVLPAVREVELYKNDVGVLEADIRF 123
DB 61 KIVR---IVDVEKVEKKFLGKPIYWKLYLEHPQDVPITREKVRHPAVVDIPEYDIPF 116
124 AMRVLIDHDLFFPTWYRVEAEPLKNGKGFYDKVYLVKSRPEPIYGERLAPTKLPDLRL 183
DB 117 AKRLIDKGL-----PNEG-----EELKIL 138
QY 184 AFDEIVSKQSGSPRRDPVIVIAKTDGDEVL-----FAEGKDEKPIREFVEY 235
DB 139 AFDEITLYHGE-BFGGPIIMISYADENKAVITWKNIDLPYEVVSSEKMKRLRI 197
QY 236 VKRYPPIIVGNNHFDWPIYLRARILGKLDVTRRGAEPPTSVMH---VSVGRIL 292
DB 198 IREKDPDIIIVYNDSPDFPLARAEGLGKLTIGRD-GSEPRQKIGDMTAVGVGRI 256
QY 293 NVLDYDAEMPEIKISLEVAEYLVGMKSERVITNMWEIPIYWDPKKRPILLQYAR 352
DB 257 HFDLHVITITINLPTTLELVYEA-LFGKPEKRYAD--EIAKAWSGEGLERVAATSM 313
QY 353 DVVATYGLAEKILPFAIOLSYVTGLPLDQVAMSVGRLEWYLRARAKKELVPRVE 412
DB 314 EDAKATYELGKPLPMEIQSLVGLGQPLMDVSRSTGNLVWELFKKAYKENGVAAPKPS 373
QY 413 RPE-----ETYGATVLEPLRGVHENIAVIDSSMTPNMIKXNVGPDILVRGEKGE 466
DB 374 EEEYORLRBSYTGKPFYKEPEKGMENIVYLDPRALYPSIITITNVPDTLNEG----- 428

QY 467 CGCMF-APVYKRRFRCPBGFKTIVLERLLKRAVRAEKKYPDPSPRYLLDEROKAL 525
DB 429 CKXNDYAPGVGHKFCXDIPIPSLLGHLLEEROKIKTKKE-TDPIEXILLIDYROKAI 487
QY 526 KVLANSYGYMGSGARWYRECAKAVTAGRLITTA-INIRKGLKTYIGDTSLFV 584
DB 488 KLANSTYGYGAKARWYKECAESTYAMGRKTYELWKELEKGFVLYIDTGLTA 547
QY 585 TYD-----PEKVENFIKIKEELG--FEIKLEKYKLLFTEAKRYVAGLLEDRIDI 635
DB 548 TIGGSESEIKKALAEVRYKINSKLPGLLEVEYGGYKGFV-TKKRAVVIDEKKVIT 606
QY 636 VGEPAVRGDMCELAKEVQTVVAVIVUKTSEVNAVEYRKIVLEBKVPYIKLYWKT 695
DB 607 RGLIEVRDMSEIAKETQARVLETILKHGVGEBAVIVKEVLOKLANYEIPPEKAIYEO 666
QY 696 LSKRLEETTEAPHVVAAKRMLSGYRVSPDGKIGVIVYKGGGRISQRAVPMYFMV--DP 753
DB 667 ITRPLEHYKAIQHVAVAKKLAAGKVKIKPGWVIGIYILRGDGPISNRA--LLAEYDP 723
QY 754 S-QIDVTYVVDHQLIPALRLIYGFITEKKLKASATGQKTLFDFL-AKSK 803
DB 724 KKAQVDAEYIENQVPAVLRILEGFGYRKEDLRVQKTDQVGLTSLWLNKKSZ 776

RESULT 7
US-07-966-278-1
Sequence 1, Application US/07966278
Patent No. 5489523
GENERAL INFORMATION:
APPLICANT: Mathur, Eric A
TITLE OF INVENTION: EXONUCLEASE-DEFICIENT THERMOSTABLE
TITLE OF INVENTION: PYROCOCUS FURIOSUS DNA POLYMERASE I
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thomas Fitting
STREET: 12526 High Bluff Road, Suite 300
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92130
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/966,278
FILING DATE: 19921226
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: STG0133P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-792-3680
TELEFAX: 619-792-8477
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-966-278-1

Query Match 29.5%; Score 1236.5; DB 1; Length 775;
Best Local Similarity 36.2%; Pred. No. 2.7e-105;

Wed Apr 23 08:09:10 2003

us-10-034-849-2.fai

Page 6

Matches 301; Conservative 153; Mismatches 283; Indels 95; Gaps 23;

```
QY 8 VDSSEYVGVKPEPOVIIMGIAENGSRVVLIDSPRPYVLLAAGADPKQV---AQRIR 63
DB 2 ILVDVITTEBEGKPVIRLFR-KENGKFKIHDHTRFPYIYALLRDSKIEEVKKTGERHG 60
QY 64 ALSRPSPIIGVEDDKRYGRRRLRIRTYLPAVAREYELVKNVGVYDLVLEADIRF 123
DB 61 KIVR---IVDEKVEKFKLPTWKLYLEHPDQVPTIREKREHRAVVDIFEVDIF 116
QY 124 AMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLVSREPLYGALAPTKLPDLRL 183
DB 117 AKRYLIDKGLI-----PMEGE-----BELKIL 138
QY 184 AFDIEVYSKQSPRPERDPVIAVKTDDGDEVL-----FLAEKQDRKPIREFVEY 235
DB 139 AFDIETLYHEGE-EFGKGPIMISYADENBAKVTIMKNIIDLPYEVVSSREMIKRFIRI 197
QY 236 VKRYDDIIVGVNNHNPMPYLLRRARILGILDTVRVGAEPPTSVHGH---VSVPGR 292
DB 198 IREKDPDIIVTNGDSDFPYLAKRAEKLGILITIGRD-GSEPMQORIGMTAVVEYKRI 256
QY 293 NVLDYVAEEMPEIKISLEEVAYGVWKKSERVIIMMDEIPYDDPKKRLLOQAR 352
DB 257 HEDLYHVITRTINLPTTYLLEAVYEA-FGKPEKRYAD--ELAKMESGENLERVAKISM 313
QY 353 DVVRATYGLAEKILIPALQISYVTGLPLDOVGAMSVGRLEWYLIRAAFMKELVNRVE 412
DB 314 EDKATYELGKFLPMEIQLSRVGLPMDVSRSSGTGMLVWFLIRAAVERNEVANKKS 373
QY 413 RPE-----ETVRCALVLEPLRGVHENIIVLDFSSMYPNIMIKYVGPDTLVVRPEKSG 466
DB 374 EEEYORRLRESYTGTFVKEPEKGLMENIYVLDFFALYPSIITINNSPDLNLEG----- 428
QY 467 CGCWE-APEVGRFRRCPPGFYTVLERLLELRKRAVMAKCPYPPSPERYLDEROKAL 525
DB 429 CKNYDIAPQVGHKCKDIPGIFPSLGHLEEROKIKTKYME-TODIEKILIDYRQKAI 487
QY 526 KYLANASGYWGMGAMWYCRECAKAVTAMGRHLRFA-IVIRAKGLKLYIGDDPSLFV 584
DB 488 KLANAFYGYGVYAKAMWYCKEACASVTAMGRKYLEWVKELBEKSGFKVLYIDDLGYLA 547
QY 585 TYD-----PEKVENFIKIKKEELG--FEIKLEKVVYRILPFTAKKRYVGLLEGRDI 635
DB 548 TIRGESEBEIKKALBEFVKYINSKLPGLELEBEFYGKGFV-TKGRIVIDEBSKVI 606
QY 636 VGFELVRGDMGELAKVQTVNVIYVKTSEVNAKVEYRKIVLEEGKVPLEKVIYMT 695
DB 607 RLEIVRDMSEINKETQAVLETLIHGDVBEAVRIVKEVYIQCLNAYEILPEPKLAIYEQ 666
QY 696 LSKRLEETAPHVAVAKMLSAGYVSPGDKIGYIVLGGGRISGRAMPYFMVK--DQ 753
DB 667 ITRPLHEYKALGPVAVAKKLAKGVKIKPGVIGIYVLRGDDGISRFA---ILAEYDP 723
QY 754 S---QIDVTYVHQIIPALRILGYFGITEKKLAKASATGQKTLFDFL-AKKS 802
DB 724 KKHKDAEYIENQVLPAVRLIRLEGYGRKEDLRYOKTRQVGLTSMINIKKS 775
```

RESULT 8
US-08-424-921-1
Sequence 1, Application US/08424921
Patent No. 5545552
GENERAL INFORMATION:
APPLICANT: Macbur, Eric A
TITLE OF INVENTION: PURIFIED THERMOSTABLE PYROCOCUS
TITLE OF INVENTION: FURTHER DNA POLYMERASE I
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bingham & Fitting
STREET: 12526 High Bluff Road, Suite 300
CITY: San Diego
STATE: CA

COUNTRY: USA
ZIP: 92130

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,921
FILING DATE: 19-APR-1995

CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/803,627
FILING DATE: 02-DEC-1991

APPLICATION NUMBER: US 07/620,568
FILING DATE: 03-DEC-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/657,073
FILING DATE: 19-FEB-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/776,552
FILING DATE: 15-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: STG0100P
TELEPHONE: 619-792-8477
TELEFAX: 619-792-3680

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYDROTHERMAL: NO
ANTI-SENSE: NO

US-08-424-921-1

Query Match 29.5%; Score 1236.5; DB 1; Length 775;
Best Local Similarity 36.2%; Pred No. 2.7e-105;
Matches 301; Conservative 153; Mismatches 283; Indels 95; Gaps 23;

```
QY 8 VDSSEYVGVKPEPOVIIMGIAENGSRVVLIDSPRPYVLLAAGADPKQV---AQRIR 63
DB 2 ILVDVITTEBEGKPVIRLFR-KENGKFKIHDHTRFPYIYALLRDSKIEEVKKTGERHG 60
QY 64 ALSRPSPIIGVEDDKRYGRRRLRIRTYLPAVAREYELVKNVGVYDLVLEADIRF 123
DB 61 KIVR---IVDEKVEKFKLPTWKLYLEHPDQVPTIREKREHRAVVDIFEVDIF 116
QY 124 AMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLVSREPLYGALAPTKLPDLRL 183
DB 117 AKRYLIDKGLI-----PMEGE-----BELKIL 138
QY 184 AFDIEVYSKQSPRPERDPVIAVKTDDGDEVL-----FLAEKQDRKPIREFVEY 235
DB 139 AFDIETLYHEGE-EFGKGPIMISYADENBAKVTIMKNIIDLPYEVVSSREMIKRFIRI 197
QY 236 VKRYDDIIVGVNNHNPMPYLLRRARILGILDTVRVGAEPPTSVHGH---VSVPGR 292
DB 198 IREKDPDIIVTNGDSDFPYLAKRAEKLGILITIGRD-GSEPMQORIGMTAVVEYKRI 256
QY 293 NVLDYVAEEMPEIKISLEEVAYGVWKKSERVIIMMDEIPYDDPKKRLLOQAR 352
DB 257 HEDLYHVITRTINLPTTYLLEAVYEA-FGKPEKRYAD--ELAKMESGENLERVAKISM 313
QY 353 DVVRATYGLAEKILIPALQISYVTGLPLDOVGAMSVGRLEWYLIRAAFMKELVNRVE 412
DB 314 EDKATYELGKFLPMEIQLSRVGLPMDVSRSSGTGMLVWFLIRAAVERNEVANKKS 373
QY 413 RPE-----ETVRCALVLEPLRGVHENIIVLDFSSMYPNIMIKYVGPDTLVVRPEKSG 466
```

```

Db 374 BEEYORRLRESYTGFGVKEPEKGLMENIYVLDPRALYPSIIITHNVSPDTLNLG----- 428
Qy 467 CGCWE-APEVKHRRFRCPGPFKTVLERLELRKRVAKEMKKYPPDSEYVLLBEROKAL 525
Db 429 CKANDIAPVOGHKFCODIPGIFPSLGLHLBEROKIKTKMKE--TODPIEKILLDYRQAI 487
Qy 526 KVLNANASYGMWGSAGRWYCEKACAVTAMGRHLIRTA-INIARKLGLKVIYGDTSLFV 584
Db 488 KLANSPFYGYGAKARWYCEKACASVTAMGRKRIELVWMLBEKGFVKLYIDTDLGYA 547
Qy 585 TYD-----PEKYENFIKIIKEBLG--FEIKLEKYVKRLFFTEAKKRYVGLLEDGRIDI 635
Db 548 TIPGSESEIKKALFEKYYINSKLPGLLELEFGFYKRGFFV--TKKRAVIDEGKVIIT 606
Qy 636 VGFEAVRGDMCELAKVOTKVEIVLKTSEVNKAVEYVRYKIVKLEBEKVILEKLVIMKT 695
Db 607 RGLIIVRDMSEIAKETQARVLETILKHGDVEAARIVKEVIOKLANEIPPEKLAIEQ 666
Qy 696 LSKRLSEYTTAPHVAAKRLMISAGYVSPGDKIGYIVKGGRISSORAMPYFMVK--DP 753
Db 667 ITRPLHEYKAIQHVAAKRLAKGVKIKRGMVIGIYVLRDGPISNRA---ILAEYDP 723
Qy 754 S--QIDVTYVVDHQQIIPALRIIGYITTEKKLNASATGQTLDPDL-AKKS 802
Db 724 KKHKYDAEYIENQVLPVAVLRILEGFGYRKEDLRYQKTRQVGLTSMNLKKS 775

```

```

RESULT 9
US-08-556-355A-1
; Sequence 1, Application US/08556355A
; Patent No. 5866395
; GENERAL INFORMATION:
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: Purified Thermostable Pyrococcus
; TITLE OF INVENTION: furiosus DNA Polymerase I
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flunegan, Henderson, Farbow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,355A
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/424,921
; FILING DATE: 19-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,627
; FILING DATE: 02-DEC-1991
; APPLICATION NUMBER:
; APPLICATION DATA:
; FILING DATE: 21-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,073
; FILING DATE: 19-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,846
; FILING DATE: 03-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. Paul
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 04121.0004-02
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-556-355A-1

```

```

Query Match 29.5%; Score 1236.5; DB 2; Length 775;
Best Local Similarity 36.2%; Pred. No. 2.7e-105;
Matches 301; Conservative 153; Mismatches 283; Indels 95; Gaps 23;

```

```

Qy 8 VLDSYEVGKEPQVLIWGIAENGERVVLIDRSFRPYFVALLAPGADPKQV---AQRIR 63
Db 2 ILVDVYTERGKPVIRLFK-KENQKFKIHDRTFRPIYVALLRDSKIEEVKKITGERHG 60
Qy 64 ALSRPKPIIGVEDDKKTKYGRPRVLRIRTVLPEAVREYRELKVQVGVEDVLEADIRF 123
Db 61 KIVR---IVDVEKVEKKFKGPIYWKLYLEHPQDVPTIREKVRBPAVDIFEYDIPF 116
Qy 124 AMRYLIDHDLFPFTWVAVAEPLNNKGFVVDKVLVKSREPPLYGBALPTKLPDLRI 183
Db 117 AKYILDKGLI-----PDEGE-----BELKIL 138
Qy 184 AFDIEVYSKQSPRPERRPVIIVAKTDGDEVL-----FLAEGKDRKPIREFVEY 235
Db 139 AFDIETLYHGE--EFGKPIIMISYADENKAVITWKNIDLPYEVVSSREMIKRLRI 197
Qy 236 VKRYPPIIYGNHNDWYLLRRARILGKLDVTRVGAEPITSHGH---VSPGRL 292
Db 198 IREKDPPIIYVNGDSDFPYLAKRAEKIGIKLTGRD-GSEPMQIRIGMTAVEVGR 256
Qy 293 NVLDYDAEEMPEIKIKISLEVAEYGVMMKSSRVLIINMEIPDYMDPKRRPLLOAR 352
Db 257 HFDLIYHITRTINLPYTLAEVTAI-FGKPKREVAD--EIAKAMSGENLERVAKYSM 313
Qy 353 DVAVATYGLAEKILPFAIQLSYVTGLPLDOVGAMSVGFLEWYLIRAFKKEIVPNRVE 412
Db 314 EDKATYELKEPLPMETIQSLRGVGLMDVSSSTGNLVEMFLRQAVRNEVAPKPS 373
Qy 413 RPE-----ETYGATVLEPLRGVHENIAVLDPSSMTPNIMIKYNGPDLVRPEKGE 466
Db 374 BEEYORRLRESYTGFGVKEPEKGLMENIYVLDPRALYPSIIITHNVSPDTLNLG----- 428
Qy 467 CGCWE-APEVKHRRFRCPGPFKTVLERLELRKRVAKEMKKYPPDSEYVLLBEROKAL 525
Db 429 CKANDIAPVOGHKFCODIPGIFPSLGLHLBEROKIKTKMKE--TODPIEKILLDYRQAI 487
Qy 526 KVLNANASYGMWGSAGRWYCEKACAVTAMGRHLIRTA-INIARKLGLKVIYGDTSLFV 584
Db 488 KLANSPFYGYGAKARWYCEKACASVTAMGRKRIELVWMLBEKGFVKLYIDTDLGYA 547
Qy 585 TYD-----PEKYENFIKIIKEBLG--FEIKLEKYVKRLFFTEAKKRYVGLLEDGRIDI 635
Db 548 TIPGSESEIKKALFEKYYINSKLPGLLELEFGFYKRGFFV--TKKRAVIDEGKVIIT 606
Qy 636 VGFEAVRGDMCELAKVOTKVEIVLKTSEVNKAVEYVRYKIVKLEBEKVILEKLVIMKT 695
Db 607 RGLIIVRDMSEIAKETQARVLETILKHGDVEAARIVKEVIOKLANEIPPEKLAIEQ 666
Qy 696 LSKRLSEYTTAPHVAAKRLMISAGYVSPGDKIGYIVKGGRISSORAMPYFMVK--DP 753
Db 667 ITRPLHEYKAIQHVAAKRLAKGVKIKRGMVIGIYVLRDGPISNRA---ILAEYDP 723
Qy 754 S--QIDVTYVVDHQQIIPALRIIGYITTEKKLNASATGQTLDPDL-AKKS 802
Db 724 KKHKYDAEYIENQVLPVAVLRILEGFGYRKEDLRYQKTRQVGLTSMNLKKS 775

```

RESULT 10

```

US-07-803-627A-1
; Sequence 1, Application US/07803627A
; Patent No. 5948663
; GENERAL INFORMATION:
; APPLICANT: MATTHEW, Eric J.
; TITLE OF INVENTION: Purified Thermotable Pyrococcus
; TITLE OF INVENTION: Purified DNA Polymerase I
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803,627A
; FILING DATE: 02-DEC-1991
; CLASSIFICATION: 415
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/776,552
; FILING DATE: 14-OCT-1991
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,073
; FILING DATE: 19-FEB-1991
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,568
; FILING DATE: 03-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. Paul
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 04121.0004-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4400
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; -07-803-627A-1
;
Query Match 29.5%; Score 1236.5; DB 2; Length 775;
Best Local Similarity 36.2%; Pred. No. 2.7e-105;
Matches 301; Conservative 153; Mismatches 283; Indels 95; Gaps 23;
;
QY 8 VLDSSYEVVKGKPOVITWIGIAENGERVVLIDSPFPYFALAGADPKOV---NQRIR 63
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2 ILVDVITEEGKPVIRLPK-KENGKFKIEHDTFRPYIYALLRDSKIEVKKITGBRG 60
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 ALSRPKSPIGVEDDKRYFGRPRVLRIRTVLPEAVREVELVANDGVEDLVADIRF 123
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 KIVR-----IVDEKVEKKEFKLQKPIWKLKYLEHPOVPIREKREHRAVVDLFEVDIRF 116
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 AMRYLLDHDLPFTWYRVEAPLENKKGFRVVDVYLVSRPEPLXGRLAPTKLPDLRL 183
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 117 AKRYLLDKGLI-----PMEGR-----EELKRL 138
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 APDIEVYSKQSPREPRDVIVAKTDGDEVL-----FLAEGKDRKRIRBPVEY 235
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 139 APDIEVLVHEGR-EFGKQPIIMISYADENEAKYITWKNIDLPYVEVVSSEKREMIKQFLRI 197
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 VKRYVDDIIVGYNNHDPWYLLRRAIIGIKLDVTRVGAEPFTSVYGH---VSPGRL 292
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 198 IREKPDIIIVYNGDSFDFPYLAERAKGKIKLTIGRD-GSPKMQRIQDMTAVEVKRI 256
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 293 NVLDYVYAEEMPEIKISLEEVAVLGVWKKSRVILNMKEIPQVDDPKRRPLLOVAR 352
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 257 HFDLYVITRTINLTPTYLEAVYPAI-FGRKPKRYAD--ELAKMSENGENLERVAKYKM 313
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 353 DVPRTATYGAELKIPPAIQLSYVTGLPFDVYGMVSGFLEWYILRAAFKKELVNRVE 412
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 314 EDKATVYELGKEPLPMEIQLSRLVGQPLDVSRSSSTGNLVEPFLRAKVERNEVAAPKFS 373
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 413 RPE-----ETRYGALVLEPLRGVHENIAYLDSSVNPYIMIKYVGPDTLVPRGKCE 466
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 374 EEEVQRRLRESYTGQFVPEPEKGLMEYIVLFRALYPSIIITHNVSBDTLNEG----- 428
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 CGCME-APVVKARFRCPGPFKTVLERLLELRKVRAMKYPDPSPREYRLDEROKAL 525
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 429 CKRYDIAPOVGHKFCOKDPCGFIPSLGLHLEEROKIKTEWKE-TODPLEKILLDYRQKAI 487
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 526 KYLANASYGMWGMGARGRYCECAKAVTAMGRHLIRTA-INIARRKGLKTYGDTSLPV 584
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 488 KLANSPFYGYVYARARVYCECAESTAMGRKYLELVWKELEKRFKVLVIDTGGVLA 547
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 585 TYD-----PEKVENFKIKIKELG--PEIKLEKRYKLEPTEAKRYAGLDERDOI 635
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 548 TITGQSESEIKKGLAEFKVINSKLPGLLELEFGYRKGFPV-TKRRVAVIDBEGVATT 606
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 636 VGPBAVRGWCGLAKEVOTKYVEIIVLKTSEVNKAVERKIVKLEBGRVPIELVIWKT 695
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 607 RGEIVRRDMSIAKETQARVLETILKHGDVEAVRIVEYIOLKLANVEIIPKRLAIYEQ 666
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 696 LSKRELEYTTEAPHVAAKRMLSAGYVSPDGKIGYIVYKGGRIISORAPYEMVK--DP 753
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 667 ITRPLHEVYKAIQPHVAVAKKILAAKGVKIKGWMYIGYIVLRGGPISRA---ILAEYDP 723
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 754 S--QIVVTYVYDQIIPALRLILGYFGITEKKLKLSATGQKTLDPFLAKKS 802
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 724 KKKYDAEYIENQVLPALVRLIILGFGYRKEDLRQYKTRQVGLTSMWIKKS 775
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;
RESULT 11
US-09-073-354-1
; Sequence 1, Application US/09073354
; Patent No. 6038659
; GENERAL INFORMATION:
; APPLICANT: KITABAYASHI, Masao
; APPLICANT: ARAKAWA, Taku
; APPLICANT: INOUE, Hiroaki
; APPLICANT: KAMAKAMI, Bunsei
; APPLICANT: KAMAMURA, Yoshihisa
; APPLICANT: IMANAKA, Tadayuki
; APPLICANT: MORIKAWA, Masaaki
; APPLICANT: TAKAGI, Masahiro
; TITLE OF INVENTION: A Thermotable DNA Polymerase and Kites for
; TITLE OF INVENTION: Amplifying Nucleic Acids
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,354
; FILING DATE: Concurrent Herewith
; CLASSIFICATION: 0506
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/656,005
; FILING DATE: 24 MAY 1996

```

APPLICATION NUMBER: JP 134096/95
 FILING DATE: 31 MAY 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Toffenetti, Judith L.
 REGISTRATION NUMBER: 39,048
 REFERENCE/DOCKET NUMBER: 2418/9
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-429-1776
 TELEFAX: 202-429-0796
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 774 amino acids
 TYPE: amino acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-073-354-1

Query Match 29.5%; Score 1234.5; DB 3; Length 774;
 Best Local Similarity 36.4%; Pred. No. 4,2e-105;

Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;

QY 8 VLDSSYEVNGKEPOVITWGAENGERVVLIDRSFRPFYALLAPGADPKQV---AQRIR 63
 2 ILTDYITDEKPVIRIFK-KENGSEFKIEYDRTEPEFYALLKDDSAIEVKKITAERHG 60
 DB 64 ALSRPKPIIGVEDDKKYPGRPRVLRIRTVLPEAVREYRELKVNVDGVEDVLEADIR 123
 61 TVVTVKR---VEKVOQKFKGRPEVWKLFTTHPDVPAIRDKIRHGAVIDIYEYDIPF 116
 QY 124 AMRYLIDHDLFPFTWYVEAEPLNKKGFVVDKYLVSREPELYGEALAPTKLPDLRL 183
 117 AKRLLIDKGLV-----PMEGD-----ELKML 138
 QY 184 AFDIEVYVSKGSPRPDPVIVIAVTKDDGSEVL-----FLAEGKDRKPIREFVEY 235
 139 AFDIQTLYHEGEBAE-GRIMISYADEEGARVITWKKVVDLPYVDVSTEREMIKRFLRV 197
 QY 236 VKRDPDIIIVGNNHPPWPLRLARLITGLDVTNRVNGEPTTSVGH---VSVPRL 292
 198 VKERDPDLITVNDNPFALYKRCERLGINFALGRD-GSEPKIQMGDRFAVEVKRI 256
 QY 293 NVLDYAEEMPEIKIKSEVAEYLVGKKSERVIIMWEIIPYWDPKRKPILLQYAR 352
 257 HFDLYPVIRINPTTYLLEAVENAVFGQPK-EKYVAE-ELTPAMEGLENLERYAKISM 313
 QY 353 DQVATYGLAEKILPFAIOLSYVGLPLDQVGVSVGRLEMYLIRAAFKMKELVPPNVE 412
 314 EDAKVTYELGKEPLPMEQQLSRLIGQSLMDVSRSGTGLVEMPLLRKAYENNELAPNKP 373
 413 RPE-----ETRGALVLEPLAGHENLAVDPSSMYPMIMIKNVGPTLVPRGKCGEC 467
 374 EKEIARRRQSYEGGVKEPERGLMENIYLDPRSLYPSIIITHNVSPPTLARE----- 426
 QY 468 GCWE---APEVGRFRRCPPGFFKTVLERLLELRKVAEMKKYPPDSPEYRLDEROKA 524
 427 GCKEYDVAPOVGHFCDPFGFISPLDGLBEKOKIKKKKA-TIDIERKLLDYRORA 485
 QY 525 LKVLANSYGMVSGARWYCRECAKAVTANGRLIRTAI-NIARKGLKAVIGDPTSLF 583
 486 IKILANSYGYGYARARWYCKECAESVTANGREYITWTIKIEIEKGFVKIYSDTDF 545
 QY 584 VTY---DPEKE---NFIKIKEL--GFEIKLEKYKLPFTPEAKRYAGLLEGRID 634
 546 ATIPGADAEIVKKAAMEFLYNALPGALEYEGYKRGPFV-TKKKAVVIDEGRKIT 604
 QY 635 IVGEAVVGDWCELAQVQTVKEIVLKTSEVNAVEVVRKIVLELGEKVPPIELGLVWK 694
 605 TRGLEIYVRDSEIAKEIQAVLEALLKQDGEVKAVRIVKYTEGLSKYEVKPEPELVH 664
 QY 695 TLSRLIEEYTTAHPVNAKMLSGAVVSPQDKIGYIVKGGGRISORAMPYFWVQPS 754
 665 QITDLDKQYKATGPHVAVAKRLAAGVXIRPGTVISYIVLKGSGRIQGRALP-FDEFPPT 723

QY 755 --QIDVTYYVDHOLIPALAILGFGITEKULKASATGKTLFDPLAKR 801
 DB 724 KHKYDAEYIENOVLPAPERILRAFGRKEDLRVQKTRQVLSAWLKP 772

RESULT 12

US-08-656-005A-1
 Sequence 1, Application US/08656005A
 Patent No. 6054301

GENERAL INFORMATION:

APPLICANT: KITABAYASHI, Masao
 APPLICANT: ARAKAWA, Taku
 APPLICANT: INOUE, Hiroaki
 APPLICANT: KAWAKAMI, Bunsei
 APPLICANT: KAWAMURA, Yoshihisa
 APPLICANT: IMANAKA, Tadayuki
 APPLICANT: TAKAGI, Masahiro
 APPLICANT: MORIKAWA, Masaaki
 TITLE OF INVENTION: A Method of Amplifying Nucleic
 TITLE OF INVENTION: Acid and A Reagent Therefor
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Kenyon & Kenyon
 STREET: 1025 Connecticut Avenue, N.W., Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Wordperfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/656,005A
 FILING DATE: 24 MAY 1996
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 134096/95
 FILING DATE: 31 MAY 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Toffenetti, Judith L.
 REGISTRATION NUMBER: 39,048
 REFERENCE/DOCKET NUMBER: 2418/3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-429-1776
 TELEFAX: 202-429-0796
 INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 774 amino acids
 TYPE: amino acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-656-005A-1

Query Match 29.5%; Score 1234.5; DB 3; Length 774;
 Best Local Similarity 36.4%; Pred. No. 4,2e-105;

Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;

QY 8 VLDSSYEVNGKEPOVITWGAENGERVVLIDRSFRPFYALLAPGADPKQV---AQRIR 63
 2 ILTDYITDEKPVIRIFK-KENGSEFKIEYDRTEPEFYALLKDDSAIEVKKITAERHG 60
 DB 64 ALSRPKPIIGVEDDKKYPGRPRVLRIRTVLPEAVREYRELKVNVDGVEDVLEADIR 123
 61 TVVTVKR---VEKVOQKFKGRPEVWKLFTTHPDVPAIRDKIRHGAVIDIYEYDIPF 116
 QY 124 AMRYLIDHDLFPFTWYVEAEPLNKKGFVVDKYLVSREPELYGEALAPTKLPDLRL 183
 117 AKRLLIDKGLV-----PMEGD-----ELKML 138

Wed Apr 23 08:09:10 2003

us-10-034-849-2.rst

Page 10

QY 184 AFDEIYVSKQSPREPERDPIVIAVKTDDGDEVL-----FLAEGDKDRKIREFVEY 235
DB 139 AFDIQTLVHGESEFAE--GPILMISYADEEGARVITWKNVDLPYDVVSTREMIKREFLV 197
QY 236 VKRYDPDIIIVGYNHNPMPYLLRARIIGIKLDVTRVGAEPSTVHGH---VSPGRL 292
DB 198 VKKEDPVLITNGDNDFPALKRCKEKGINFALGRD--GSEPKIORGDRFAVEYKGR 256
QY 293 NVDLXYAEBMEPEIKISLEVAEYLGVMKSSRVINMMWELPDVMDPKKRLLOVAR 352
DB 257 HFDLYVIRRTINLPYTLLEAVYAVGQPK--EKVYAE--ETPAMEGENSELRYARISM 313
QY 353 DDVRAITYGLAEKILFPALQLSYVTGLPLDOYGAMSVGRLELYILRAAFMKELVNRVE 412
DB 314 EDKAVTYELGKEFLPMKQSLRGLGSLMDVSSSTGNLVEMFLRKAYERNEBLAPKFD 373
QY 413 RPE-----ETRGAIVLBPLRGVHENIAVLDESSMYPNIMIKYVGPDTLVDPGEKCGEC 467
DB 374 EKEIARRROSIEGKYVEKERGMENIVLDFRSLPSTIITHNVSPDTINRE----- 426
QY 468 GCWE---ABEVKGRFRRCPPGFVTVLERLLERKRVRAEMKYPSPDSPEYRLDEROKA 524
DB 427 GCKEYDVAPOVGHRCXDPGPFIPSLIGDLLEBRQIKKMKKA--TIDPIERKLLDYROBA 485
QY 525 LKVLNASTGYMGMSGARWYCRECAKAVTAMGHLIRTAI--NIARKLGLKVIYGDPTSLF 583
DB 486 IKILANSYGYGYVARAWYCKEACSVTAMGREYITMTIKIEBKRGPFVITSDTDGFF 545
QY 584 VTY---DPEKVE---NFIKIKKEEL--GPEIKLEKVVYKRLPFTBAKKRYAGLLEGRD 634
DB 546 ATIPGADAEIVKCKMEFLNINAKLPGALELEBEFYRGFFV--TKKTAIVADEBKIT 604
QY 635 IVGFENVRDGCBLAKEVQTVNVEIVLKTSEVNAKAVEYRKIVKLEBGKVPLEKVIYK 694
DB 605 TRGLEIVRDMSEIKAEQARVALLKODGVKAVRIKVEYTEKLSYVEPPEKVIHNE 664
QY 695 TLSKRLSEYTTAPVNAKMLSAGYRVSPPDKIGYVIVKGGRISSORAMPYFMWKDS 754
DB 665 QITRIDKDYKATGPVNAVAKKLAAGYKIRPGTVISYIVLKSGSGIDDRAP--FDFEPDPT 723
QY 755 --QIDVTYVVDHIIIPALRIITLGYFTTEKKLKASATGQKTLFPLAK 801
DB 724 KHKYDAEYIENQVLPAVERILRAFGRKEDLRYQKTRQVGLSAMLKPK 772

RESULT 13

US-09-073-259-1

Sequence 1, Application US/09073259

Patent No. 6143536

GENERAL INFORMATION:

APPLICANT: IMANKA, Tadayuki

APPLICANT: TAKAGI, Masahito

APPLICANT: MORIKAWA, Masaki

TITLE OF INVENTION: DNA encoding a Thermocable DNA Polymerase

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: 1025 Connecticut Avenue, N.W., Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect 6.1 Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/073,259

FILING DATE: Concurrent Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/656,005

Query Match 29.5%; Score 1234.5; DB 4; Length 774;
Best Local Similarity 36.4%; Pred. No. 4,2e-105;
Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;
FILING DATE: 24 MAY 1996
APPLICATION NUMBER: JP 134096/95
FILING DATE: 31 MAY 1995
ATTORNEY/AGENT INFORMATION:
NAME: Tofenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 2418/10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 amino acids
TYPE: amino acid
STRANDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-259-1

QY 8 VLDSYEVVGEKPEVITINGIAENGERVVLIDRSFRPYFALLAGADPKOV---AORIR 63
DB 2 ILDTDITDGDQPVIRIK--KENGEFKIEYDRFEBFYFALLKODSAIEEVKKTIERHG 60
QY 64 ALSRPSPILIGVEDDKRYGPRPRVLRITVLPKAYRELYKANDVGEDVLEADIRF 123
DB 61 TVTVYKRV---VEKQKFLORPYEWKLTFTHPQVPAIRKDKIRHCAVIDIYTDIPF 116
QY 124 AMRYLIDHOLFPTTVRYVAEPLBNKKGFRVDKVIKVSREPELYGELAPTKLPDLRL 183
DB 117 AKRLLIDKGLV-----PMEGD-----EELKML 138
QY 184 AFDEIYVSKQSPREPERDPIVIAVKTDDGDEVL-----FLAEGDKDRKIREFVEY 235
DB 139 AFDIQTLVHGESEFAE--GPILMISYADEEGARVITWKNVDLPYDVVSTREMIKREFLV 197
QY 236 VKRYDPDIIIVGYNHNPMPYLLRARIIGIKLDVTRVGAEPSTVHGH---VSPGRL 292
DB 198 VKKEDPVLITNGDNDFPALKRCKEKGINFALGRD--GSEPKIORGDRFAVEYKGR 256
QY 293 NVDLXYAEBMEPEIKISLEVAEYLGVMKSSRVINMMWELPDVMDPKKRLLOVAR 352
DB 257 HFDLYVIRRTINLPYTLLEAVYAVGQPK--EKVYAE--ETPAMEGENSELRYARISM 313
QY 353 DDVRAITYGLAEKILFPALQLSYVTGLPLDOYGAMSVGRLELYILRAAFMKELVNRVE 412
DB 314 EDKAVTYELGKEFLPMKQSLRGLGSLMDVSSSTGNLVEMFLRKAYERNEBLAPKFD 373
QY 413 RPE-----ETRGAIVLBPLRGVHENIAVLDESSMYPNIMIKYVGPDTLVDPGEKCGEC 467
DB 374 EKEIARRROSIEGKYVEKERGMENIVLDFRSLPSTIITHNVSPDTINRE----- 426
QY 468 GCWE---ABEVKGRFRRCPPGFVTVLERLLERKRVRAEMKYPSPDSPEYRLDEROKA 524
DB 427 GCKEYDVAPOVGHRCXDPGPFIPSLIGDLLEBRQIKKMKKA--TIDPIERKLLDYROBA 485
QY 525 LKVLNASTGYMGMSGARWYCRECAKAVTAMGHLIRTAI--NIARKLGLKVIYGDPTSLF 583
DB 486 IKILANSYGYGYVARAWYCKEACSVTAMGREYITMTIKIEBKRGPFVITSDTDGFF 545
QY 584 VTY---DPEKVE---NFIKIKKEEL--GPEIKLEKVVYKRLPFTBAKKRYAGLLEGRD 634
DB 546 ATIPGADAEIVKCKMEFLNINAKLPGALELEBEFYRGFFV--TKKTAIVADEBKIT 604
QY 635 IVGFENVRDGCBLAKEVQTVNVEIVLKTSEVNAKAVEYRKIVKLEBGKVPLEKVIYK 694
DB 605 TRGLEIVRDMSEIKAEQARVALLKODGVKAVRIKVEYTEKLSYVEPPEKVIHNE 664
QY 695 TLSKRLSEYTTAPVNAKMLSAGYRVSPPDKIGYVIVKGGRISSORAMPYFMWKDS 754

Db 665 QITRDLKDYKATGPHVAARLAAAGVAKIRPGTVISYVLKSGRIGDRAIP-FDEFDP 723

Qy 755 -QIDVTYVVDHQQIIPAAALRILGFGITEKKLKSATGQKTLFDFLAKK 801

Db 724 KHKYDAEYIENQVLPAAERILRAFGYRKEDLRYOKTRQVLSAMLKPK 772

RESULT 14

US-09-363-095-1

Sequence 1, Application US/09363095

Patent No. 6187573

GENERAL INFORMATION:

APPLICANT: IMANAKA, Tadayuki

APPLICANT: TAKAGI, Masahiro

APPLICANT: MORIKAWA, Masaaki

TITLE OF INVENTION: DNA Encoding A Thermostable DNA Polymerase

FILE REFERENCE: 2418/11

CURRENT APPLICATION NUMBER: US/09/363,095

CURRENT FILING DATE: 1999-07-30

EARLIER APPLICATION NUMBER: US 09/073,259

EARLIER FILING DATE: 1998-05-06

EARLIER APPLICATION NUMBER: US 08/656,005

EARLIER FILING DATE: 1996-05-24

EARLIER APPLICATION NUMBER: JP 134096/95

EARLIER FILING DATE: 1995-05-31

NUMBER OF SEQ ID NOS: 16

SOFTWARE: WordPerfect 6.1 windows

SEQ ID NO 1

LENGTH: 774

TYPE: PRT

ORGANISM: Hyperthermophilic archaeon

US-09-363-095-1

Query Match 29.5%; Score 1234.5; DB 4; Length 774;

Best Local Similarity 36.4%; Pred. No. 4.2e-105;

Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;

Qy 8 VLDSEYVNGKEPOVITMGANGERVYLIDRSFRPYFALLAGADPKQY----AQRIR 63

Db 2 ILDDYITDEGKQVIRIRFK-KENGEFKIEYRTTEPYFALLKQDSAEVKKITAEBSHG 60

Qy 64 ALSRPKSPILGVEDDKRYGPRPRVLRIRTVLPEAVREYRELKVNQVGVDELADIRF 123

Db 61 TVTVVKKR---VEKVKQKFKGRFVEMKLYFTHPQDVPAIDKIREHNAVIDIYEYDIPF 116

Qy 124 AMRLIHDHLPFTWYRVEAPLENKGFVQDKYLVKSRPEPLYGEALAPTKLPDLRIL 183

Db 117 AKRYLIDKGLV-----PMEGD-----EELKML 138

Qy 184 AFDEIVYSGKSPRRERDPVIVIAVKTDGDEVL-----FLAEGDDKRPREFEY 235

Db 139 AFDIQTLYHGESEFAR-GPILMISYADSEGARVITKXNDLPYADVSTEREMKRFPLRV 197

Qy 236 VKRYDPIIVGVNNHNPWLYLRARILGKLDVTRVGALEPTTSVGH---VSPGRL 292

Db 198 VKKQDPVLIYVNDGNDFPAILKKRCGLINPALGR-GSEPKIQRGKDFPAVEVGRIR 256

Qy 293 NVLDIYDAEMPEIKISLEVAEVLGVKKSERVIINMWEIPYWDPKKRPLLOYAR 352

Db 257 HFDLYPYRRTINLPYTLAEVAVFQPK-EKYVAE--EITPAMEGENILERVARYSM 313

Qy 353 DDVATYGLAEKILPFAIOQSVYVGLPLDVGAMSVGRFLMYYLIRAFKKEVLPRNVE 412

Db 314 EDAVYTELGEPLPMEQSLRLGGOSIMDVSRSTGALVWFFLRKAYENNELAPNKP 373

Qy 413 RPE-----ETRYGAILVEPLRGVHENIADVLFSSVYPMIMKYVGPPTLVPRPGKSEC 467

Db 374 EKELARRQSYEGGVYKPERGMLNENIYVLDRLSLYPSIIITTHVSPPTLARE----- 426

Qy 468 GCWE---APEVKAHFRRCPPGFFKTVLERLLLELRKRVAEKKYPPDSEYRLDEROKA 524

Db 427 GCKEYDVAPOVGHFRFCDFPGFISLLGLDLLEEROKIKKKMKA-TIIDIERKLLDYRQRA 485

Qy 525 LKVLNANSGYMGWMSGARWYCRECAKAVTAMGRHLIRTAI-NIARKLGLKVIYGDTSLF 583

Db 486 IKILANSYGYIGYARARWCKECSAVTAMGREYITMTKEIEKYGFFVITSDTGGF 545

Qy 584 VTY---DPEKVE---NFIKIIEEL--GFEIKLEKYYKELFTEAKRYAGLLEDGRID 634

Db 546 ATIPGDAEIVKKKAMEFLVYINAKLPGLALELIEYEGYKRGFFV-TKKKAVYIDEBGKIT 604

Qy 635 IVGFVAVRGWCLAKEVQTRVVEIVLKTSEVNKAVEYKRIYKELEBEGVPIELKVIWK 694

Db 605 TRGLEIVRRDMSIETAKETQARVLEALLKQDVKEKAVIYKEVTEKLSKYEVPEKLVIEH 664

Qy 695 TLSKRLEETTEAPHVAAAKRMLSAGRVSPGKIGVVIYKGGRIISQRAWPFYFWDPDS 754

Db 665 QITRDLKDYKATGPHVAARLAAAGVAKIRPGTVISYVLKSGRIGDRAIP-FDEFDP 723

Qy 755 -QIDVTYVVDHQQIIPAAALRILGFGITEKKLKSATGQKTLFDFLAKK 801

Db 724 KHKYDAEYIENQVLPAAERILRAFGYRKEDLRYOKTRQVLSAMLKPK 772

RESULT 15

US-09-418-027-1

Sequence 1, Application US/09418027

Patent No. 6225065

GENERAL INFORMATION:

APPLICANT: KITABAYASHI, Masao

APPLICANT: ARAKAWA, Taku

APPLICANT: INOUE, Hiroaki

APPLICANT: KAWAKAMI, Bunsei

APPLICANT: KAWAMURA, Yoshihisa

APPLICANT: IMANAKA, Tadayuki

APPLICANT: TAKAGI, Masahiro

APPLICANT: MORIKAWA, Masaaki

TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for

TITLE OF INVENTION: Amplifying Nucleic Acids

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: 1025 Connecticut Avenue, N.W., Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect 6.1 windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/418,027

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/073,354

FILING DATE:

APPLICATION NUMBER: JP 134096/95

FILING DATE: 31 MAY 1995

ATTORNEY/AGENT INFORMATION:

NAME: Tofteneft, Judith L.

REGISTRATION NUMBER: 39,048

REFERENCE/DOCKET NUMBER: 2418/9

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-429-1776

TELEFAX: 202-429-0796

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 774 amino acids

TYPE: amino acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-418-027-1

Wed Apr 23 08:09:10 2003

us-10-034-849-2.ra1

Page 12

Query Match 29.5%; Score 1234.5; DB 4; Length 774;
Best Local Similarity 36.4%; Pred No. 4,26-105;
Matches 302; Conservative 145; Mismatches 289; Indels 93; Gaps 22;

```
QY 8 VIDSSEYVGEKEPOVITWGAENGERVVIDSFRPYFALAPGADPKOV---AQRIR 63
DB 2 IIDDYITDEGKAVIRIFK-KENGEFKIRYDSEBYFALKODSAIEVKKITTHERHG 60
QY 64 ALSRPKSPITIGVEDDKKRTFGRRVLRIRITVLPFAVREYRELVKQVDEYDLADIRF 123
DB 61 TVVTVKR---VEKVOKKFLGRPEVEWKLYFTHPQDVPAIRDKIREHGAVIDIYEYDIPF 116
QY 124 AMRYLIDHDLFPFTWYRVEAEPLNKGFRVDKYLKSRPEPLYGEALAPTKLPDLRL 183
DB 117 AKRYLIDKGLV-----PMEGD-----BELKML 138
QY 184 AFDEIVYSKQSPRPREDPVIVIAVKTDDGEVL-----FIAEGKDRKPIREPEY 235
DB 139 AFDIQTLYHEGEFPAE-GPILMISYADEEGARVITWKVVDLPYDVVSTEREMIKRFLRV 197
QY 236 VRRYDPOIIVGNNHDFMYPILRRARILIGLIDVTRVGAEPSTVHG--VSVPGR 292
DB 198 VKENDPDLITINDNDNFAYLKRCCKGKGINFALGRD-GSEPKIQMGDRFAVEYKRI 256
QY 293 NVLDYDAEEMPEIKISLEEVAVYGVWKKSEYIINWWEIPDYDDPKRPLIQYAR 352
DB 257 HFDLPVIRRTINLPYITTEAVTEAVFGQPK-EKVYAE-EITPAMETGENLERVARISM 313
QY 353 DDVATYGLAEKILPFAIQLSYVTGLPLDQVGMASVGRLEWYLIRAAFRKELVPRNVE 412
DB 314 EDKATYELGKEFLPMEQOLSRILIGSLMDVSRSTGNLVEMFLRQAYERNELAPNKPD 373
QY 413 RPE---ETVGAIVLEPLRGVHENIAVLDPSSMYPNIMIKYVGPDTLVRPGEKGCBC 467
DB 374 EKELARRROSYESGVEKPERGLMENIVYLDPRSLYPSIITHNVSPTLNR----- 426
QY 468 GCME---APEVGRFRRCPPGFKTYLERLTELRRVRAEMKKYPPSPPEYRLDEROKA 524
DB 427 GCEYDVAPQVGRFCFCDPGFIPSLIGDLLEBOKIKKKWKA-TIDPIERKLDYRQRA 485
QY 525 LKVLNANASYGMGSGARWYCRCAKAVTAMGRHLIRPAI-NIARKLGLKVIYGDTSLF 583
DB 486 IKILANSYGYGYARARWYCKRCASVTAMGREYITMTIKEIBEKYGFRVYISDTDGF 545
QY 584 VTY---DPEKVE---NEIKIIEEL--GEIKLEKYKRLFTFEAKRYAGLLEBRID 634
DB 546 ATIPGADAEIVKKGAMFLLNYINAKLPGLALEIEYGFYKRGFV-TKKYIAVIDEGRKIT 604
QY 635 IVGFPAVAGDWCELAKEVQTKVEIVLKTSEVNKAVEYRKIVKELEBGRVPLEKLVIRK 694
DB 605 TRGLEIVARDMSEIATKETAQARVEALLLKDGDEKAVRIKVEVTEKLSKEYEVPPEKLVIE 664
QY 695 TLSKRLSEYTTAEHPVVAANKRMLSAGYRVSFGDKIGYIVYVKGSGRISORAMPYFWKDP 754
DB 665 QITRDLKDYKATGPHVAARLARAGVKIRPGTVISYIVLKGSGRIDRAIP-FDEPDP 723
QY 755 --QIDVTYVDHDIIPALALRIIGFSITTEKLLKASATGOKTLFDFLAKK 801
DB 724 KHKYDAEYIENQVLPAVERILRAFGYRKEDLRQKTRQVGLSAMLKPK 772
```

Search completed: April 23, 2003, 07:31:47
Job time : 33 secs